

Generalized Tikhonov Regularization: Topological Aspects and Necessary Conditions

Von der
Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von
Nadja Sofia Worliczek
geboren am 17.08.1978
in Wien/Österreich

Eingereicht am: 17.12.2013

Disputation am: 31.03.2014

1. Referent: Prof. Dr. Dirk Lorenz

2. Referent: Prof. Dr. Thorsten Hohage

2014

Contents

Contents	iii
Introduction	v
1 Similarity and convergence issues	1
1.1 Preliminaries on topological spaces	1
1.2 Preliminaries on sequential convergence spaces	5
1.3 Prametrics	11
1.3.1 The prametric topology	12
1.3.2 Sequential convergence structures induced by prametrics	17
2 Ill-posed inverse problems and regularization	23
2.1 Inverse problems and terms of solution	23
2.2 Well-posedness and stability	29
2.3 Regularization	33
3 Generalized Tikhonov regularization	39
3.1 Basic notation	39
3.2 Formalization of Tikhonov setups	40
3.3 Some necessary conditions on regularizing Tikhonov setups . .	43
4 Necessary conditions on a common set of sufficient conditions	47
4.1 A common set of sufficient conditions	47
4.2 CONV and CONT	50
4.2.1 Bottom slice topologies	51
4.2.2 Necessary conditions	55
5 The special case of Bregman discrepancies	59
5.1 Bregman distances - definition and basic properties	60
5.2 Bregman prametrics, topological issues and Tikhonov regularization	66

Conclusion	73
Notation	75
Index	78
Bibliography	81
Zusammenfassung (Summary in German)	87

Introduction

Context, motivation and subject

An inverse problem is the problem of retrieving some quantity of interest from some measured data, that is in some way indirectly related to our quantity of interest, e.g. by some physical process. Prominent examples are computed tomography, inverse scattering problems and impedance tomography.

Mathematically, an inverse problem is usually modelled in the following way: denote by X the set of quantities of interest, i.e. of the objects we are originally interested in, and by Y the set of observable data.

This allows us to describe the relation between observed data and quantities of interest by a mapping

$$F : X \rightarrow Y$$

which maps the quantity of interest to the actual measurement. Now the inverse problem consists simply in solving an operator equation of the form

$$Fx = y.$$

While this sounds great in theory, in reality we are confronted with some obstacles, from the mathematical as well as from the practical point of view. First, the operator F may not be injective. Also, in reality, the right hand side is obtained from some sort of measurement process and measurements are usually noisy. So the data may not even be contained in the image of F and therefore the equation may not have a solution respectively we may not obtain the 'exact' solution.

Commonly we are not able to avoid noise completely, so what we wish at least, is that the solution for data 'near' to exact data, is, given it exists, near to the exact solution and if we approach exact data, the solutions approach the exact solution.

This is reflected by the idea of well-posedness, which can be traced back to Jaques Hadamard ([Had02]), i.e. a problem is well-posed if it has a solution for every right hand side, every such solution is unique and the solutions

depend in some sense 'continuously' on the data. Clearly, this concept applies to every category that allows a notion of 'nearness', 'getting near' and 'continuity'.

Unfortunately, most real world inverse problems are ill-posed. Nevertheless, we want to solve them, so we have to find a workaround. A first step in this direction is to redefine the term solution, e.g. by taking the Moore-Penrose inverse or minimum norm solutions, which is mainly an attempt to cure the lack of existence and uniqueness of a solution.

In common, this approach alone is not enough to transform our problem to a stable one. Additionally, the problem has to be regularized. The idea of regularization consists essentially in replacing the original inverse problem \mathcal{P} with a series of 'similar' problems that approach the original problem in some way in dependence of some parameter.

There are various definitions of what is understood exactly under the term of 'regularization' in several particular kinds of settings. Since we aim at analysing a whole class of regularization methods itself, in particular on necessary conditions on the interplay of the involved objects to form such a regularization, we are in need for an axiomatic definition ourselves. After revising and discussing some of these definitions and concepts, namely from [EHN96], [Rie03] and [SKHK12], such an axiomatization will be given in Chapter 2, intended to merge the essential ideas for a range of settings as broad as possible.

Before we are able to do so, we have to answer some questions, namely:

- How do we describe 'nearness', 'similarity' and 'getting near'?
- What does 'solving an inverse problem' actually mean?
- What are we hoping to achieve by regularizing an inverse problem?
- What do we actually understand by stability and ill- respectively well-posedness?

Concerning the first question, the classical answer is: 'consider neighbourhoods, distance and convergence in some class of metric spaces'. While the idea of taking general metric spaces is rather old, see for example [TA77], the probably best studied particular class of metric spaces are Hilbert spaces (e.g. [Eng81], [CK94],[GHS11a]). In the last decade, the focus shifted to general Banach spaces (e.g. [BO04], [HKPS07]).

Also, already since the earlier days of regularization theory, the restriction to metric spaces was felt to be rather unsatisfactory, be it due to practical or due to theoretic reasons. Therefore general topological spaces were taken

into consideration (e.g. by [Iva69], [GHS11b]). Unfortunately, they usually lack some of the main features of metric spaces which make the latter more easily accessible in the context of regularization of inverse problems. So in general, nearness in the sense of smallness of neighbourhoods can not be described (and consequently not be compared) by means of real numbers, nor can convergence of sequences be characterized by convergence of real sequences, which deprives us of the well-developed toolkits of optimization and real analysis.

Perhaps even more important, from the modeller's point of view, it would sometimes be desirable to use specific non-metric bivariate real valued functionals to measure nearness because they reflect the statistical nature of noise, the particular way in which noise is applied or other essential aspects of the given problem particularly well, while they blind out distinguishing features which are not of importance (see e.g. [Pö08], [Fle10], [Fle11], [BB11] and [HW13]). Unfortunately, there does not necessarily exist a topology such that getting small of such a functional, i.e. getting near, is equivalent to convergence.

Moreover, regularization methods are meant to be translated to algorithms that can be handled by computers and therefore we are mainly interested in 'getting near' in the sense of convergence of sequences, which can often be obtained by less restrictive assumptions than would be required for convergence in terms of filters or nets.

On the whole, these points impose the question whether the category of topological spaces is really the right category to deal with regularization of inverse problems. Starting from these considerations, in Chapter 1, we will present the category of sequential convergence spaces which might be a possible alternative. Amongst other things, the question, when sequences are enough to describe a topology and relations between sequential convergence spaces will be discussed. As pointed out, describing nearness and convergence in terms of real numbers is a very convenient feature. So, we make also allowance for this and study topologies and sequential convergence structures induced by so called prametrics, that is, extended real valued functionals that meet the minimum requirement of vanishing on the diagonal of their domain of definition. Afterwards, the other items of the above list will be discussed in Chapter 2.

A very widely used class of regularization methods consists in the class of (generalized) Tikhonov regularization methods, where the surrogate problems consist in minimizing functionals of the form

$$\rho(Fx, y) + \alpha R(x),$$

where ρ is some 'measure of nearness' called discrepancy functional. There is

a vast number of publications dealing with that sort of regularization methods, again mostly using Hilbert space settings and powers of norms as discrepancy functionals. Therein, the choice of R and convergence rates were the main subjects of interest (see e.g. [CK94], [DDD04]). Lately, generalizations to more general Banach spaces and to non-metric discrepancy functionals and related topologies are increasingly studied, e.g. in [AV94], [BO04], [Res05], [RA07], [HKPS07], [Pö08], [Fle10] and [Fle11]. In general they use sets of conditions on the interplay of the involved components, being sufficient to ensure that the inverse problem under question is regularized in some way, and most of them containing variants of standard assumptions which allow to apply standard techniques from calculus of variations as e.g. the direct method. This raised the following questions to us, which will be the subject of Chapter 3 and Chapter 4, namely:

- Are these standard assumptions necessary and which mix of ingredients has a chance to actually end up in a regularization method?
- Are there interdependences between these assumptions?
- When do we get along with a purely topological setting, i.e. a setting where all convergences are induced by topologies?

Structure of this thesis

Chapter 1 The first chapter is dedicated to the study of similarity and convergence issues. After recalling some facts from the theory of topological spaces, we introduce the category of sequential convergence spaces plus some of its basic properties. Furthermore, we investigate some topologies and sequential convergence structures related to prametrics.

Chapter 2 In the second chapter we develop a possible formal definition for regularization methods, which provides the opportunity to work with mixed settings of topological spaces and sequential convergence spaces. In preparation for this task, we discuss what 'solution of an inverse problem', 'well-posedness' and related terms are supposed to mean and settle on formal definitions suitable enough to work with in the further course.

Chapter 3 The third chapter deals with generalized Tikhonov regularization. It starts with the basic idea and notation and goes ahead by defining terms called variational setup and regularizing Tikhonov setup, the first being essentially a list of the mathematical objects involved in

Tikhonov regularization and the second being a variational setup from which a regularization method for the (also contained) inverse problem can be formed. It is finally ended by some necessary conditions for a variational setup to be a regularizing Tikhonov setup.

Chapter 4 In the fourth chapter, we discuss a common set of conditions sufficient for a variational setup to be a regularizing Tikhonov setup and take a closer look on the interdependence of some of its components. In particular, we are interested in the question, when two of this assertions can be fulfilled simultaneously by a purely topological setting and end up by necessary conditions for such a case in terms of bottom slice topologies, which are again topologies related to prametrics and are also of interest in their own right.

Chapter 5 In the last chapter finally, some of the terminology developed in the preceding chapters is applied to the special case of Bregman discrepancies. We construct prametrics based on Bregman distances, take a closer look on some of their properties and sequential convergence with respect to corresponding bottom slice topologies and end up with a regularizing Tikhonov setup involving such topologies.

Acknowledgements

In the first place, I want to thank my advisor Dirk Lorenz for being a friendly, helpful, motivating and and at times demanding mentor, and Thorsten Hohage for taking the time to act as second reviewer for this thesis.

In addition, I want to thank Jens Flemming and Christiane Pöschl for sending me their doctoral theses and all the reviewers of [LW13] for a lot of fruitful questions and comments, for finding errors, for pointing out things that needed more explanation and finally for reminding me that thought reading isn't a very common virtue and hence shouldn't be demanded.

Finally, special thanks to all the persons, who supported me on the mainly non-mathematical level, in particular to Christoph Köhler and to the 'Kaffeerunde' for its steady supply of coffee, cake and humour.

Chapter 1

Similarity and convergence issues

1.1 Preliminaries on topological spaces

For the convenience of the reader we recall some definitions and some more or less well-known facts from topology which will be used in various places of this thesis.

We will be particularly interested in sequential properties of topologies and in the question, when topologies can be completely described by such sequential properties. A nice survey on that subject, which was extensively used for this section and served as a formidable signpost to literature on that issue, can be found in [Gor04].

The purpose of the first definition is simply to allow for a more compact formulation of a frequently used condition. It does not depend on topological spaces and will also be used in other contexts.

Definition 1.1.1

Let X be a set and $U \subseteq X$. Then a sequence $(x_n)_{n \in \mathbb{N}}$ is said to be *eventually in U* if there is an index n_0 such that $x_n \in U$ for all $n \geq n_0$.

Now we would like to recall some frequently used terms concerning sequences in topological spaces.

Definition 1.1.2

Let (X, τ) be a topological space.

- (i) A subset $U \subseteq X$ is called *sequentially open*, if every sequence $(x_n)_{n \in \mathbb{N}}$ in X converging to an element $x \in U$ is eventually in U .

- (ii) A subset $U \subseteq X$ is called *sequentially closed*, if the limits of every convergent sequence in U belong to U .
- (iii) A subset $U \subseteq X$ is called *sequentially compact*, if every sequence in U has a subsequence converging to an element of U .
- (iv) The topology τ is called sequential if all sequentially open sets are open or equivalently all sequentially closed sets are closed.

We want to emphasize the well-known fact, that sequentially open sets behave in its relation to sequentially closed sets as would be expected from open sets in relation to their closed counterpart.

Remark 1.1.3

A set is sequentially open if and only if its complement is sequentially closed and vice versa.

There are some prominent examples as well of sequential topologies as of non-sequential ones.

Example 1.1.4

- (i) All metric topologies are sequential.
- (ii) Weak topologies on infinite dimensional locally convex spaces are in general not sequential, e.g. the weak topology on ℓ^1 is not sequential.

Sequential topologies provide some interesting features. One of them is, that continuity of a mapping starting from a sequential space can completely be characterized by sequences.

Lemma 1.1.5 (Continuity w.r.t sequential topologies)

Let X be equipped with a sequential topology τ_X , (Y, τ_Y) an arbitrary topological space and $F : X \rightarrow Y$ a mapping.

Then F is continuous if and only if it is sequentially continuous.

Proof: See e.g [Gor04]. □

In Chapter 4 it will be useful to know if subspaces of sequential spaces are sequential. In general, this is not the case as has been shown in [Fra65]. Nevertheless, there are cases where this property is passed on to a subspace.

Proposition 1.1.6 ([Fra65])

Let (X, τ) be a sequential topological space and $M \subseteq X$ be open or closed. Then the subspace topology $\tau|_M$ on M induced by τ is sequential.

Proof: Since the proof is omitted in the cited paper, we give a short sketch. For the case of an open set the subspace topology on M is given by

$$\tau|_M = \{U \subseteq M \mid U \in \tau\}$$

and since τ is sequential, it is sufficient to show that every $U \subseteq M$ which is sequentially open with respect to $\tau|_M$ is sequentially open with respect to τ . Now M is itself sequentially open, hence we can assume without loss of generality that every sequence with τ -limit in U lies completely in M . Now every τ -limit in U of such a sequence is also a limit with respect to $\tau|_M$, and since U is sequentially open with respect to $\tau|_M$, the sequence is eventually in U which proves the claim.

For M closed, the closed sets with respect to $\tau|_M$ are exactly the τ -closed subsets of M . Let U be a $\tau|_M$ -sequentially closed set. Because M is τ -closed the τ -limits of every sequence in U are in M , hence the sequence converges also with respect to $\tau|_M$ and therefore every such limit is in U . \square

On the other hand, we can in general not hope that arbitrary subspaces are sequential, for this does only hold for an even more special class of topological spaces.

Lemma 1.1.7 ([Fra67])

Let (X, τ) be a sequential topological space. Then the following statements are equivalent:

- (i) All topological subspaces of (X, τ) are sequential.
- (ii) (X, τ) is a Fréchet space, i.e. the closure of a set is the set of the limits of all sequences in that set.

In Chapter 4 we will study some topologies constructed as initial topologies. Therefore we also recapitulate their definition.

Definition 1.1.8

Let X be a set, $((Y_i, \tau_i))_{i \in I}$ a family of topological spaces and denote by $\mathcal{F} := (f_i : X \rightarrow Y_i)_{i \in I}$ a family of mappings.

The coarsest topology on X with respect to whom all the f_i , $i \in I$ are continuous is called *initial topology* with respect to \mathcal{F} .

Initial topologies can be explicitly constructed as follows.

Lemma 1.1.9 ([Bou66, Ch. I, §2, Prop. 4])

Let X , (Y_i, τ_i) and \mathcal{F} be as above. Then the set

$$\mathcal{B} := \{f_i^{-1}(U_i) \mid U_i \in \tau_i, i \in I\}$$

is a subbase for the initial topology with respect to \mathcal{F} , i.e. a set is open in the initial topology if and only if it is a union of finite intersections of elements of \mathcal{B} .

In the context of initial topologies we are interested in sequential convergence as well, which can be characterized as below.

Lemma 1.1.10

Let X , (Y_i, τ_i) and \mathcal{F} be as above.

A sequence (x_n) in X converges to $x \in X$ in the initial topology with respect to \mathcal{F} if and only if $f_i(x_n) \xrightarrow{\tau_i} f_i(x)$ for all $i \in I$.

Proof: Sequential continuity of continuous mappings yields the first direction. The converse is an immediate consequence of Lemma 1.1.9. \square

In Chapter 4 we will consider initial topologies induced by functional mapping to $[0, \infty]$ equipped with the order topology (with respect to ' $<$ '). Therefore, we recollect some facts about this topology in the following lemma.

Lemma 1.1.11 (The order topology on $[0, \infty]$)

- (i) The order topology on $[0, \infty]$ is the topology generated by all sets of the form

$$[0, b[,]a, b[\text{ and }]a, \infty], a, b \in [0, \infty].$$

- (ii) The open rays

$$[0, b[\text{ and }]a, \infty], a, b \in [0, \infty]$$

constitute a subbase of the order topology on $[0, \infty]$.

- (iii) The order topology is the one-point compactification of $[0, \infty[$ equipped with the standard topology.

Proof:

- (i) See e.g. [Mun00, Chap. 2, §14].
- (ii) Is an immediate consequence of (i), since every open interval is a finite intersection of such open rays.
- (iii) $[0, \infty]$ is a compact Hausdorff space. Furthermore the set $[0, \infty[$ is dense in $[0, \infty]$ and equipped with the standard topology is a proper subspace of $[0, \infty]$. Since $[0, \infty] \setminus [0, \infty[$ contains only a single point, the definition of one-point compactification given e.g. in [Mun00, Chap. 3, §29] applies.

\square

1.2 Preliminaries on sequential convergence spaces

In the introductory chapter we asked the question, if topological spaces are always the appropriate category to handle inverse problems. This question was based on the point of view, that we are mainly interested in getting near to exact solutions if we get near to exact data. Also it is common practice to restrict oneself to sequential convergence and other sequential analogues to various topological terms (see e.g. [HKPS07], [Fle10], [HW13]).

In this section we introduce the category of sequential convergence spaces, which we will propose as, at least partial, alternative to topological spaces later on. Sequential convergence spaces are, as topological spaces, a possibility to generalize metric spaces, starting from the point of a priori given convergent sequences as can be done for metric spaces. In the case of unique limits, they are also known as L-spaces and are rather old, according to [BB02] and [Kis60] going back to Fréchet ([Fre06]) and being already studied by Urysohn ([Ury26]).

We will give some basic definitions and some facts concerning the relation of sequential convergence spaces to topological spaces and illustrate them by some examples. For more details we refer to the textbook [BB02], in which most of the given results were found and which also discusses the more general construction of convergence spaces that allows to carry over some fundamental results of functional analysis.

Definition 1.2.1 (Sequential convergence structure)

- (i) A mapping

$$\mathcal{S} : X \rightarrow 2^{\{\text{sequences in } X\}}$$

is a *sequential convergence structure* on a set X if the following two axioms are fulfilled:

(S1) For each $x \in X$ the constant sequence $(x_n)_{n \in \mathbb{N}}$ given by $x_n = x$ is contained in $\mathcal{S}(x)$.

(S2) If $(x_n)_{n \in \mathbb{N}} \in \mathcal{S}(x)$ then so does every subsequence of $(x_n)_{n \in \mathbb{N}}$.

- (ii) We call a pair (X, \mathcal{S}) *sequential convergence space*, if \mathcal{S} is a sequential convergence structure on X .

- (iii) If $(x_n)_{n \in \mathbb{N}} \in \mathcal{S}(x)$, $x \in X$ does hold, we say that $(x_n)_{n \in \mathbb{N}}$ *converges (with respect to \mathcal{S}) to x* , write $x_n \xrightarrow{\mathcal{S}} x$ and call x an \mathcal{S} -*limit* of $(x_n)_{n \in \mathbb{N}}$.

As in every category, we want to be able to study relations between sequential convergence spaces by looking at their morphisms. Moreover, the aim of this thesis is to use sequential convergence spaces to examine approximate solvability of operator equations, which will be impossible if we do not establish a connection between mappings and sequential convergence spaces.

Definition 1.2.2

Let (X, \mathcal{S}_X) and (Y, \mathcal{S}_Y) be two sequential convergence spaces. A mapping $F : X \rightarrow Y$ is called \mathcal{S}_X - \mathcal{S}_Y -continuous, if $(x_n)_{n \in \mathbb{N}} \in \mathcal{S}_X(x)$ implies that $(F(x_n))_{n \in \mathbb{N}} \in \mathcal{S}_Y(F(x))$ for all $x \in X$.

By $\mathcal{C}_s(X, Y)$ we denote the set of \mathcal{S}_X - \mathcal{S}_Y -continuous mappings from X to Y .

Amongst other things, the notion of continuity allows us, to compare different convergence structures on the same set.

Definition 1.2.3

Let \mathcal{S}_1 and \mathcal{S}_2 be two sequential convergence structures on a set X .

We call \mathcal{S}_1 *stronger* than \mathcal{S}_2 (and \mathcal{S}_2 *weaker* than \mathcal{S}_1) if the identity mapping is \mathcal{S}_1 - \mathcal{S}_2 -continuous or equivalently $x_n \xrightarrow{\mathcal{S}_1} x$ implies $x_n \xrightarrow{\mathcal{S}_2} x$ for all sequences (x_n) in X .

After having nearly drowned the reader in a flood of definitions, we are keen to animate them a little bit by some examples.

Example 1.2.4 (Sequential convergence structure of a topology)

Let τ be a topology on X . Then $\mathcal{S}(\tau)$ given by $(x_n) \in \mathcal{S}(\tau)(x)$ if and only if $x_n \xrightarrow{\tau} x$ is a sequential convergence structure on X . We call it the *sequential convergence structure induced by τ* .

In the case of two topological spaces (X, τ_X) and (Y, τ_Y) and a mapping $F : X \rightarrow Y$, $\mathcal{S}(\tau_X)$ - $\mathcal{S}(\tau_Y)$ -continuity turns out to be the same as the well-known concept of sequential continuity.

Two examples of such convergence structures are weak convergence on a normed space X and weak*-convergence on its dual space X^* , the first is induced by the weak topology on X and the latter by the weak*-topology on X^* . Since they are initial topologies with respect to $(x \mapsto x^*(x))_{x^* \in X^*}$ respectively with respect to $(x^* \mapsto x^*(x))_{x \in X}$ this is a consequence of Lemma 1.1.10.

Example 1.2.4 also allows us to define a meaningful term of continuity for mappings starting from a topological space and mapping to a sequential convergence space. This will allow us to mix categories in the context of regularization and consequently to make terminology applicable to a broader range of settings.

Definition 1.2.5

Let (X, τ) be a topological space and (Y, \mathcal{S}) be a sequential convergence space. Then a mapping $F : X \rightarrow Y$ is called τ - \mathcal{S} -continuous if it is $\mathcal{S}(\tau)$ - \mathcal{S} -continuous. Conversely, a mapping $G : Y \rightarrow X$ is called \mathcal{S} - τ -continuous if it is \mathcal{S} - $\mathcal{S}(\tau)$ -continuous.

Now, we get back to examples of sequential convergence spaces.

Example 1.2.6

- (i) Let X be a set of mappings from a set M to an arbitrary topological space. Then pointwise convergence is a sequential convergence structure on X .
- (ii) Let (Ω, Σ, μ) be a measure space and X be a set of measurable real valued functions on Ω . Then convergence almost everywhere is a sequential convergence structure on X .
- (iii) A further class of examples, which will be discussed in detail in Section 1.3, is given by sequential convergence structures induced by prametrics, see Definitions 1.3.1 and 1.3.10 for details.

One of the conditions which are demanded by Seidman in his definition of an approximation scheme given in [Sei81] is convergence of a sequence of functions in an 'appropriate sense', which is for the case of Efimov-Stečkin space specified as so called graph subconvergence. This subsequently presented example shows, that sequential convergence spaces are still too restrictive in the sense, that there are convergences which do not define a sequential convergence space but are nevertheless suitable for speaking of approximation in a sensible way.

Example 1.2.7 (Subconvergence of sets)

In [Sei81] notions of subconvergence of sets and graph subconvergence of mappings are defined in a special setting, which can nearly literally be carried over to arbitrary topologies.

So, let (X, τ) be a sequential convergence space. We call a sequence (M_n) in the power set 2^X of X *subconvergent* to $M \in 2^X$, if the τ -limits of every convergent subsequence (x_{n_k}) such that $x_{n_k} \in M_{n_k}$ are contained in M .

Clearly, every subsequence of a subconvergent sequence of sets in 2^X is again subconvergent to the same set, since it already contains all limits of arbitrary subsequences.

But a constant sequence with single member M subconverges to M if and only if the limits of any convergent sequence in M are in M , i.e. M is sequentially closed. So, subconvergence defines a sequential convergence structure on a subset $\mathfrak{M} \subset 2^X$ if and only if all members of \mathfrak{M} are sequentially closed.

A sequence $(f_n)_{n \in \mathbb{N}}$ of mappings $f_n : D_n \subseteq X \rightarrow Y$ into another topological space (Y, τ_Y) is called *graph subconvergent* to $f : D \subseteq X \rightarrow Y$ on $X \times Y$ if the sequence of graphs $\Gamma_n := \{(x, f_n(x)) \mid x \in D_n\}$ subconverges to the graph Γ of f in $X \times Y$ endowed with the product topology. As mentioned also in [Sei81], this applies to a constant sequence if and only if its only member has sequentially closed graph. Seidman also stresses, that the notion of graph subconvergence applies to set-valued mappings as well.

Since all definitions rely only on sequential convergence, they can be easily translated to sequential convergence spaces instead of topologies.

The last example already shows, that, unfortunately, not every notion of convergence can be covered by terms of sequential convergence structures. Another case of 'convergence' that is not given by a sequential convergence structure and is closely related to subconvergence, is subsequential convergence. We define it nevertheless because it is frequently used as appropriate notion of convergence to define stability and convergence of regularization methods.

Example 1.2.8

Let X be endowed with a sequential convergence structure or a topology respectively. A sequence $(x_n)_{n \in \mathbb{N}}$ is said to *converge subsequentially* to $x \in X$ if (x_n) has a subsequence converging to x (with respect to the sequential convergence structure or the topology). Every limit of a subsequence of (x_n) is called a *subsequential limit*.

Subsequential convergence is in general not a sequential convergence structure, because the axiom (S2) is usually violated as one sees e.g. for the sequence (x_n) in \mathbb{R} equipped with the standard topology given by

$$x_n = \begin{cases} n & \text{for } n \text{ even} \\ \frac{1}{n} & \text{for } n \text{ odd,} \end{cases}$$

which converges subsequentially to zero but has a subsequence that does not. Moreover, if a subsequential convergence coming from a topology would fulfil axiom (S2), then convergence of one single subsequence would be enough to show convergence of the whole sequence with respect to the defining topology, due to the Urysohn-property (see Definition 1.2.14 below).

Not very surprisingly, in Example 1.2.4 we have seen that every topology induces a sequential convergence structure, and one is tempted to ask, if maybe every sequential structure is given in this way. As it happens, the answer is no, see Remark 1.2.12 for a counterexample. This circumstance gives rise to the following definition.

Definition 1.2.9

- (i) A sequential convergence structure \mathcal{S} on a set X is called *topological* if there exists a topology τ on X such that $\mathcal{S} = \mathcal{S}(\tau)$.
- (ii) A sequential convergence space (X, \mathcal{S}) is called *topological* if \mathcal{S} is topological.

As we will see below, not every sequential convergence structure is topological. However, it is possible to construct a topology from an arbitrary sequential convergence structures \mathcal{S} , such that all \mathcal{S} -convergent sequences are also convergent with respect to the topology. For this purpose we introduce some more terminology, again in perfect analogy to known terms from topology.

Definition 1.2.10

Let (X, \mathcal{S}) be a sequential convergence space.

- (i) A set $U \subseteq X$ is called *sc-open* with respect to \mathcal{S} if every \mathcal{S} -convergent sequence (x_n) with limit in U is eventually in U .
- (ii) A point $x \in X$ is called *\mathcal{S} -limit point* of $U \subseteq X$ if there exists a sequence (x_n) in $U \setminus \{x\}$ such that $x_n \xrightarrow{\mathcal{S}} x$.
- (iii) We denote the set of sc-open subsets by $\tau(\mathcal{S})$ and call it the *topology induced by \mathcal{S}* .

Now we have to justify the last definition and take the opportunity to present some further properties of topologies induced by sequential convergence structures taken from [BB02] and [Kis60].

Lemma 1.2.11

Let (X, \mathcal{S}) be a sequential convergence space.

- (i) The set $\tau(\mathcal{S})$ is indeed a topology on X .
- (ii) A set A is closed with respect to $\tau(\mathcal{S})$ if and only if it contains all its \mathcal{S} -limit points.
- (iii) If \mathcal{S} provides unique limits, then $(X, \tau(\mathcal{S}))$ is a T_1 -space, i.e for any $x_1 \neq x_2 \in Y$ there are open neighbourhoods $U(x_1)$ and $U(x_2)$ such that $x_1 \notin U(x_2)$ and $x_2 \notin U(x_1)$.

Proof:

(i) Clearly \emptyset and X are in $\tau(\mathcal{S})$. Let $U = U_1 \cap U_2$ where $U_1, U_2 \in \tau(\mathcal{S})$ and $x_n \xrightarrow{\mathcal{S}} u \in U$. Since U_1 and U_2 are sc-open, (x_n) is eventually in both sets and therefore it is eventually in U . If U is the union of an arbitrary number of sc-open sets, then every member u of U lies at least in one of these sets and so does the tail of every sequence \mathcal{S} -converging to u from some index on.

(ii) Let A be closed. If there would exist a sequence in A converging to $x \in X \setminus A$ it would have a subsequence in $Y \setminus A$ since $Y \setminus A$ is open. This is impossible.

Now let A be a set which contains all its \mathcal{S} -limit points. If $X \setminus A$ would not be open there would exist a sequence \mathcal{S} -converging to $x \in X \setminus A$ which has a subsequence in A which would be again a contradiction.

(iii) E.g. $U(x_1) = X \setminus \{x_2\}$, $U(x_2) = X \setminus \{x_1\}$, see [Kis60].

□

Now as promised, we give an example of a non-topological sequential convergence space.

Remark 1.2.12 ([Ord66])

Let (Ω, Σ, μ) a measure space with σ -finite measure μ . Then convergence almost everywhere on a set X of measurable real-valued functions is in general not topological:

If there is a sequence (f_n) in X which converges in measure to $f \in X$, then σ -finiteness of μ implies, that every subsequence of (f_n) has a subsequence converging a.e. to f (see e.g. [Bau68, 19.6]). Since (f_n) does not converge a.e. itself, this convergence can not be topological.

Such a sequence exists e.g. in the set of bounded Lebesgue-measurable functions on $\Omega = [0, 1]$.

Even in the case of a topological sequential convergence space, the inducing topology is not uniquely defined, as the following example shows.

Example 1.2.13

Due to Schur's Lemma a sequence in ℓ^1 is weakly convergent if and only if it converges with respect to the norm topology. But since ℓ^1 has infinite dimension, the norm topology and the weak topology are distinct.

A popular way to show convergence of a sequence in topological spaces, is to deduce it from convergence of all subsequences of subsequences. Again, in sequential convergence spaces, this useful tool is in general not available, so disposability of this technique deserves its own name.

Definition 1.2.14 (Urysohn-property)

A sequential convergence space (X, \mathcal{S}) is said to have the *Urysohn-property* if a sequence $(x_n)_{n \in \mathbb{N}}$ converges to $x \in X$ provided every subsequence of (x_n) has a subsequence converging to x .

Besides of its technical advantages, the Urysohn-property provides a necessary condition for sequential convergence spaces to be topological (already used in Remark 1.2.12) and in the case of unique limits it is even characteristic for topologicality.

Lemma 1.2.15 (see [BB02, Prop. 1.7.15], [Kis60])

Let (X, \mathcal{S}) be a sequential convergence space.

- (i) If \mathcal{S} is topological, then (X, \mathcal{S}) has the Urysohn property.
- (ii) If \mathcal{S} has unique limits and the Urysohn property, then \mathcal{S} is induced by a topology. Notably $\mathcal{S}(\tau(\mathcal{S})) = \mathcal{S}$ does hold.

1.3 Prametrics

As pointed out in the introduction, in the context of inverse problems it is desirable to express 'similarity' of two elements of a given set Y by means of real numbers respectively of a bivariate functional that serves as similarity measure, i.e. small values indicate similarity whereas big values stand for dissimilarity. While, as also stressed in the introductory chapter, metrics are sometimes too restrictive and it may be convenient to skip symmetry and triangle inequality, it is reasonable to demand that such a similarity measure respects equality, i.e. it vanishes on the diagonal of the Cartesian square.

This chapter will be dedicated to the study of functionals that meet this minimum requirement and of structure they establish on a set.

Following [AF90] we will address this class of functionals as follows.

Definition 1.3.1

Let Y be a set. We call a mapping $\rho : Y \times Y \rightarrow [0, \infty]$ a *prametric* if $\rho(y, y) = 0$ does hold for all $y \in Y$.

We call a prametric ρ on Y *separating* if $\rho(y_1, y_2) = 0$ implies $y_1 = y_2$.

Remark 1.3.2

Prametrics are also referred to as premetrics in several places and the term used by us seems to appear rarely in literature. Nevertheless, we decided to follow the English translation [AF90] of the Russian original because, in contrast to the term of premetric, it provides the inestimable advantage of not being used for different functionals of a similar kind too.

Remark 1.3.3

Note, that we consider the functional $(y_1, y_2) \mapsto \tilde{\rho}(y_1, y_2) := \rho(y_2, y_1)$ as different from ρ . Due to the possible lack of symmetry the functional ρ may exhibit completely different behaviour in each of its arguments, heavily influencing the outcome of some constructions based on prametrics which will be given later on in this book.

Example 1.3.4

- (i) Clearly every metric on Y is a prametric.
- (ii) Bregman distances (see Definition 5.1.1) restricted to the domain of the subdifferential of their inducing functional are prametrics, and so are Bregman prametrics, which are prametrics of such restricted Bregman distances to a possibly larger set (see also Definition 5.1.1). In particular the Kullback-Leibler divergence (see Example 5.1.3) restricted to an appropriate set is a prametric.
- (iii) The functional $\text{eq} : Y \times Y \rightarrow [0, \infty]$ given by

$$\text{eq}(y_1, y_2) = \begin{cases} 1 & \text{if } y_1 \neq y_2 \\ 0 & \text{else} \end{cases}$$

is a prametric.

- (iv) Let Y and Z be (possibly distinct) sets and $S : Y \times Z \rightarrow [0, \infty]$. Then the mapping $S_Y : Y \times Y \rightarrow [0, \infty]$ given by

$$S_Y(y_1, y_2) := \inf_{z \in Z} (S(y_1, z) + S(y_2, z))$$

is a prametric on Y if and only if for every $y \in Y$ there exists a sequence $(z_n) \in Z$ such that $\lim_{n \rightarrow \infty} S(y, z_n) = 0$.

This is e.g. the case if for all $y \in Y$ the set $\{z \in Z \mid S(y, z) = 0\}$ is non-empty.

This functional is taken from [Fle11, Def. 2.7], where it is used as an instrument for comparing elements of Y despite the similarity measure he uses for Tikhonov regularization is defined on $Y \times Z$.

1.3.1 The prametric topology

Knowing the metric topology, it seems inviting to construct topologies from arbitrary prametrics in a perfectly analogous way. This approach works indeed and we get topologies that exhibit some nice properties, i.e. even

if we loose in general most of the advantages of metric topologies as the first axiom of countability (see [AF90, 2.4]), they can at least be completely described by means of sequences.

Implementing this construction (also to be found in [AF90]) leads to the following (for the present merely formal) definition.

Definition 1.3.5

Let (Y, ρ) be a prametric space.

- (i) For $\varepsilon > 0$ and $y_0 \in Y$, we call the set

$$\mathcal{B}_\varepsilon^\rho(y_0) := \{y \in Y \mid \rho(y_0, y) < \varepsilon\}$$

the ε -ball with respect to ρ centered at y_0 .

- (ii) We call the set $\tau_\rho \subseteq 2^Y$ given by

$$\tau_\rho := \{U \subseteq Y \mid \forall y \in U \exists \varepsilon > 0 \text{ such that } \mathcal{B}_\varepsilon^\rho(y) \subseteq U\}$$

the *prametric topology* induced by ρ on Y .

The term 'topology' is justified by the following statement.

Theorem 1.3.6 ([AF90])

Let ρ be a prametric on Y and τ_ρ as in definition 1.3.5. Then the following assertions hold:

- (i) A set $A \subseteq Y$ is of the form $A = Y \setminus U$ with $U \in \tau_\rho$ if and only if $\rho(u, A) := \inf_{a \in A} \rho(u, a) > 0$ for all $u \in Y \setminus A$.
- (ii) τ_ρ is a topology on Y , i.e. the elements of τ_ρ satisfy the axioms for the open sets of a topology.

Proof: Since the proofs of both statements are omitted in [AF90], we do them ourself.

- (i) Let $U \in \tau_\rho$, $A := Y \setminus U$ and $u \in Y \setminus A = U$. Then there exists an $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon^\rho(u) \subseteq U$. Since $A \cap U = \emptyset$, the inequality $\rho(u, a) \geq \varepsilon > 0$ does hold for all $a \in A$ and therefore $\rho(u, A) > 0$.

Conversely, let $\rho(u, A) > 0$ hold for all $u \in U := Y \setminus A$. For arbitrary $u \in U$ set $\varepsilon(u) := \frac{\rho(u, A)}{2}$. Then for all $\tilde{u} \in \mathcal{B}_{\varepsilon(u)}^\rho(u)$ the inequalities

$$\rho(u, \tilde{u}) < \varepsilon(u) < \rho(u, A) = \inf_{a \in A} \rho(u, a)$$

do hold and hence $\tilde{u} \notin A$. Therefore we have $\mathcal{B}_{\varepsilon(u)}^\rho(u) \subseteq Y \setminus A = U$ and consequently $U \in \tau_\rho$.

- (ii) It is sufficient to show, that the sets of the form $Y \setminus U$, $U \in \tau_\rho$ satisfy the axioms for the closed sets of a topology.

Clearly \emptyset and Y are in τ_ρ and therefore the sets $\emptyset = Y \setminus Y$ and $Y = Y \setminus \emptyset$ can be represented as desired.

Now consider A_1, \dots, A_n where $A_k = Y \setminus U_k$ with $U_k \in \tau_\rho$ for all $k = 1, \dots, n$. We have to show, that there exists an $U \in \tau_\rho$ such that $A := \bigcup_{k=1}^n A_k = Y \setminus U$.

Let $y \in Y \setminus A$. Because of $Y \setminus A = \bigcap_{k=1}^n Y \setminus A_k$ and part (i) we get

$$\rho(y, A) = \inf_{a \in A} \rho(y, a) = \min\{\rho(y, A_1), \dots, \rho(y, A_n)\} > 0.$$

Since y was taken arbitrary, this implies $A = Y \setminus U$ for some $U \in \tau_\rho$.

Finally, let $A_k = Y \setminus U_k$ with $U_k \in \tau_\rho$ and $k \in I$ with an arbitrary index set I .

Then

$$U := Y \setminus \bigcap_{k \in I} A_k = \bigcup_{k \in I} Y \setminus A_k = \bigcup_{k \in I} U_k.$$

does hold. Therefore, for arbitrary $u \in U$ there exists a $k \in I$, such that $u \in U_k$. Since $U_k \in \tau_\rho$ there exists $\varepsilon > 0$, such that $\mathcal{B}_\varepsilon^\rho(u) \subseteq U_k \subseteq U$ and hence $U \in \tau_\rho$. So, $\bigcap_{k \in I} A_k = Y \setminus U$ with $U \in \tau_\rho$.

□

Two of the convenient features of metric spaces are, that first, the topology can completely be described by its convergent sequences and second, that convergence of sequences is equivalent to convergence of certain real sequences, namely of the corresponding image sequences under the metric. While only one direction of the latter stays true, the first property can be transferred to the general prametric case.

Lemma 1.3.7

Let ρ be a prametric on a set Y . Then the following statements hold:

- (i) Convergence of $\rho(y, y_n) \rightarrow 0$ implies $y_n \xrightarrow{\tau_\rho} y$.
- (ii) The topology τ_ρ is sequential. In particular, mappings to an arbitrary topological space are τ_ρ -continuous if and only if they are sequentially continuous.

Proof:

-
- (i) Let (y_n) be a sequence such that $\rho(y, y_n) \rightarrow 0$ and $U \in \tau_\rho$ an open neighbourhood of y .

Then there exists an $\varepsilon > 0$ such that $y \in \mathcal{B}_\varepsilon^\rho(y) \subset U$. Since $\rho(y, y_n)$ converges to zero $\rho(y, y_n) < \varepsilon$ for all $n \in \mathbb{N}$ large enough, and so (y_n) is eventually in $\mathcal{B}_\varepsilon^\rho(y) \subset U$. Therefore y_n converges to y with respect to τ_ρ .

- (ii) This proof is in essence a written out version of the sketch of proof given for [AF90, §2, Prop. 9].

Since all closed sets of an arbitrary topology are also sequentially closed, it is sufficient to show, that all sequentially closed sets are closed.

So, let $A \subseteq Y$ be a sequentially closed set. Consider $y \in Y$ such that $0 = \rho(y, A) = \inf_{a \in A} \rho(y, a)$. Then there is a sequence (a_n) in A , such that $\rho(y, a_n) \rightarrow 0$. Due to (i) $a_n \xrightarrow{\tau_\rho} y$ does hold and since A is sequentially closed, this implies $y \in A$. Consequently, we have $\rho(y, A) > 0$ for all $y \notin A$. Hence Theorem 1.3.6 shows that A is closed.

□

As mentioned earlier, there are fundamental differences between the special case of a metric and general prametrics.

Remark 1.3.8

- (i) The ε -balls $\mathcal{B}_\varepsilon^\rho(y_0)$ are in general neither themselves open nor is y_0 necessarily an inner point of ε -balls $\mathcal{B}_\varepsilon^\rho(y_0)$. In particular, they are in general not a neighbourhood basis for y_0 .

Consider the following example, which has already been mentioned in [LW13, Ex. 3.10]. Let $Y = M^2$ where M is an arbitrary set and let $\rho : Y \times Y \rightarrow [0, \infty]$ be the prametric given by

$$\rho(y, \tilde{y}) = \begin{cases} 0, & \text{if } \#\{i \mid i = 1, 2, \tilde{y}_i = y_i\} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

for $y = (y_1, y_2)$ and $\tilde{y} = (\tilde{y}_1, \tilde{y}_2)$. Now let $\varepsilon > 0$. The ε -balls centered at $y = (y_1, y_2) \in Y$ take the form

$$\mathcal{B}_\varepsilon^\rho(y) = \begin{cases} \mathcal{B}_1^\rho(y) = (\{y_1\} \times M) \cup (M \times \{y_2\}), & \text{if } \varepsilon \leq 1 \\ Y, & \text{otherwise} \end{cases}.$$

Since a set $U \subseteq Y$ is open if and only for each $u_0 \in U$ there is an $\varepsilon > 0$ such that the ε -ball centered at u_0 is a subset of U we get $\tau_\rho = \{\emptyset, Y\}$, i.e. the trivial topology.

- (ii) Some necessary conditions for the converse of 1.3.7 (i), will be discussed in detail in Subsection 1.3.2, see Lemma 1.3.16 and the following. In general, this is not true. Taking ρ from item (i), we get a topology τ_ρ with respect to which every sequence is converging to every $y \in Y$. In particular, every sequence chosen from $Y \setminus \mathcal{B}_1^\rho(y)$ converges to y . But we have $\rho(y, y_n) \rightarrow 0$ if and only if $(y_n)_{n \in \mathbb{N}}$ is eventually in $\mathcal{B}_1^\rho(y) \subsetneq Y$.
- (iii) A sequence may have multiple limits, see e.g. Example 1.3.18.

Since the analysis carried out in Chapter 4 will be in general not applicable to the whole space Y , we are also interested in the subspace topology induced by a prametric topology.

Clearly, the restriction of a prametric ρ to an arbitrary subset $M \subset Y$ is again a prametric. This raises the question, how the subspace topology on M induced by the prametric topology τ_ρ is related to the prametric topology induced by the restriction of ρ to $M \times M$. A partial answer is given in the following lemma.

Lemma 1.3.9 (Subspaces and prametric topologies)

Let ρ be a prametric on Y and $M \subseteq Y$. Then the following assertions are true:

- (i) $(\tau_\rho)|_M \subseteq \tau_{(\rho|_{M \times M})}$
- (ii) If M is τ_ρ -open, then $(\tau_\rho)|_M = \tau_{(\rho|_{M \times M})}$ does hold.
- (iii) If M is τ_ρ -closed, then $(\tau_\rho)|_M = \tau_{(\rho|_{M \times M})}$ does hold.

Proof:

- (i) Let $U \in (\tau_\rho)|_M$. Then there is a set $W \in \tau_\rho$ such that $U = W \cap M$. Now consider $u_0 \in U$ arbitrary. Since $u_0 \in W$, there exists $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon^\rho(u_0) \subseteq W$. Moreover,

$$\mathcal{B}_\varepsilon^{\rho|_{M \times M}}(u_0) = \{m \in M \mid \rho(u_0, m) < \varepsilon\} = \mathcal{B}_\varepsilon^\rho(u_0) \cap M \subseteq W \cap M = U$$

does hold and hence $U \in \tau_{(\rho|_{M \times M})}$.

- (ii) We only have to show $\tau_{(\rho|_{M \times M})} \subseteq (\tau_\rho)|_M$.

Let $U \subseteq M$ be an $\tau_{(\rho|_{M \times M})}$ -open set. Since the elements of $(\tau_\rho)|_M$ are exactly the τ_ρ -open subsets of M it suffices to show that for every $u \in U$ there exists an $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon^\rho(u) \subseteq U$.

Let u be in $U \subseteq M$. Since M is τ_ρ -open there exists an ε_1 , such that $\mathcal{B}_{\varepsilon_1}^\rho(u) \subseteq M$ and since U is $\tau_{(\rho|_{M \times M})}$ -open there exists ε_2 such that

$\mathcal{B}_{\varepsilon_2}^{\rho_{|M \times M}}(u) \subseteq U$. Set $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$. Because of $\varepsilon \leq \varepsilon_1$ we have $(u, y) \in M \times M$ for all $y \in \mathcal{B}_{\varepsilon}^{\rho}(u)$. This implies

$$\rho_{|M \times M}(u, y) = \rho(u, y) < \varepsilon \leq \varepsilon_2$$

and consequently $y \in U$ for all $y \in \mathcal{B}_{\varepsilon}^{\rho}(u)$.

- (iii) Since M is closed, it is sufficient to show that every $\tau_{(\rho_{|M \times M})}$ -closed subset U of M is also τ_{ρ} -closed.

Let $U \subseteq M$ be $\tau_{(\rho_{|M \times M})}$ -closed and $y \in Y \setminus U$. If $y \in Y \setminus M$, Theorem 1.3.6 yields

$$\rho(y, U) = \inf_{a \in U} \rho(y, a) \geq \inf_{m \in M} \rho(y, m) = \rho(y, M) > 0.$$

Otherwise we have $y \in M \setminus U$ and since U is $\tau_{(\rho_{|M \times M})}$ -closed, we get

$$\rho(y, U) = \rho_{|M \times M}(y, U) > 0.$$

In both cases, applying Theorem 1.3.6 completes the proof. □

1.3.2 Sequential convergence structures induced by prametrics

As already mentioned, one of the advantages of metric spaces is, that convergence of a sequence can be completely characterized by convergence of real sequences given by the metric, i.e. a sequence (y_n) in a metric space (Y, μ) converges to $y \in Y$ if and only if $\mu(y, y_n)$ tends to zero. In other words, the sequential convergence structure induced by the metric topology is exactly given in this way.

Now we are interested in the question, how these two sequential convergence structures are related to each other for general prametrics.

For our inquiries upon this subject, the following notion will be convenient:

Definition 1.3.10

Let ρ be a prametric on Y . Then we call the mapping

$$\mathcal{S}_{\rho} : Y \rightarrow 2^{\{\text{sequences in } Y\}}$$

$$y \mapsto \left\{ (y_n)_{n \in \mathbb{N}} \mid y_n \in Y, \quad \lim_{n \rightarrow \infty} \rho(y, y_n) = 0 \right\}$$

the *sequential convergence structure induced by ρ* .

This definition is justified by the remark below.

Lemma 1.3.11

Let ρ be a prametric on Y . Then \mathcal{S}_ρ is a sequential convergence structure on Y . Moreover, \mathcal{S}_ρ has the Urysohn-Property.

Proof: We have to check the axioms from definition 1.2.1. Since $\rho(y, y) = 0$ for all $y \in Y$ condition (S1) does hold. Now let $y \in Y$ and $(y_n)_{n \in \mathbb{N}} \in \mathcal{S}_\rho$ and $(y_{n_k})_{k \in \mathbb{N}}$ a subsequence of (y_n) . Then $(\rho(y, y_{n_k}))_{k \in \mathbb{N}}$ is a subsequence of the real sequence $(\rho(y, y_n))_{n \in \mathbb{N}}$ and consequently converges to the same limit. Hence $(y_{n_k}) \in \mathcal{S}_\rho(y)$ does hold.

Now we prove the Urysohn-property. Let (y_n) be a sequence in Y , such that every subsequence has a subsequence ξ converging to $y \in Y$, i.e. $\rho(y, \xi) \rightarrow 0$ for every such subsequence. Since \mathbb{R} has the Urysohn-property, this implies $\rho(y, y_n) \rightarrow 0$ and therefore $y_n \xrightarrow{\mathcal{S}_\rho} y$. \square

Remark 1.3.12

Using Definition 1.3.10 the assertion of Lemma 1.1.5 (i) (i.e. $\rho(y, y_n) \rightarrow 0$ implies $y_n \xrightarrow{\tau_\rho} y$) reads as follows: \mathcal{S}_ρ is stronger than $\mathcal{S}(\tau_\rho)$.

Now we attend to the question, when \mathcal{S}_ρ is topological and derive some necessary conditions on a topology to induce \mathcal{S}_ρ . This question is of interest in the context of inverse problems, since a positive answer is part of a standard set of assumptions (stated in Section 4.1) which is sufficient to show regularization properties of Tikhonov functionals with discrepancy term given by ρ . As we will see, the prametric topology plays a special role amongst the candidates for such an inducing topology, if, of course, there is any at all. In Chapter 4 this distinguished position of τ_ρ will be used as a tool to analyse the interdependence of some of the conditions from the mentioned set of assumptions.

Theorem 1.3.13 (see also [LW13])

The following assertions hold:

- (i) Let τ be a topology on Y . Then the sequential convergence structure $\mathcal{S}(\tau)$ induced by τ is weaker than \mathcal{S}_ρ if and only if τ is coarser than τ_ρ .
- (ii) If \mathcal{S}_ρ is topological, then $\mathcal{S}_\rho = \mathcal{S}(\tau_\rho)$. In particular, τ_ρ is the finest topology with that property.

Proof:

-
- (i) Let τ be a topology such that $\mathcal{S}(\tau)$ is weaker than \mathcal{S}_ρ . Suppose there does exist an $U \in \tau$ that is not contained in τ_ρ .

Then there is an element $u \in U$ such that $\mathcal{B}_{\frac{1}{n}}^\rho(u) \setminus U$ is non-empty for all $n \in \mathbb{N}$. Consequently we can choose a sequence (y_n) with $y_n \in \mathcal{B}_{\frac{1}{n}}^\rho(u) \setminus U$ for all $n \in \mathbb{N}$. Evidently $\rho(u, y_n) \rightarrow 0$ does hold. Since $\mathcal{S}(\tau)$ is weaker than \mathcal{S}_ρ this implies $y_n \xrightarrow{\tau} u$ in contradiction to $y_n \notin U$ for all $n \in \mathbb{N}$.

Conversely, let τ be a topology that is coarser than τ_ρ . Then every τ_ρ convergent sequence is also τ -convergent, and due to Lemma 1.3.7 (i) the assertion follows.

- (ii) Let τ be a topology on Y such that $\mathcal{S}(\tau) = \mathcal{S}_\rho$ and (y_n) a τ_ρ -convergent sequence with limit y . Due to (i) τ is coarser than τ_ρ , therefore $y_n \xrightarrow{\tau} y$ and hence $\rho(y, y_n) \rightarrow 0$.

□

Besides providing us a necessary condition for \mathcal{S}_ρ being topological, Theorem 1.3.13 is very convenient for proving some structural statements about the prametric topology itself. First we get another representation of τ_ρ .

Lemma 1.3.14

The topologies τ_ρ and $\tau(\mathcal{S}_\rho)$ coincide.

Proof: Since $\tau(\mathcal{S}_\rho)$ -convergence is weaker than \mathcal{S}_ρ we have $\tau(\mathcal{S}_\rho) \subseteq \tau_\rho$.

For the converse direction let $U \subseteq Y$ be τ_ρ -closed. Since τ_ρ -convergence is weaker than \mathcal{S}_ρ every \mathcal{S}_ρ -limit point of U is also a τ_ρ -limit of a sequence in U . Now, U as a closed set of the sequential topology τ_ρ is sequentially closed and therefore contains all the limits of all its sequences. In particular it contains all its \mathcal{S}_ρ -limit points and due to Lemma 1.2.11 it is also $\tau(\mathcal{S}_\rho)$ -closed. □

Second, the result about the subspace topology induced by the prametric topology given in Lemma 1.3.9 is now merely an immediate consequence of Theorem 1.3.13.

Corollary 1.3.15

Let ρ be a prametric on a set Y and $M \subseteq Y$. Then the subspace topology $(\tau_\rho)|_M$ on M is coarser than the prametric topology induced by the restriction $\rho|_{M \times M}$.

We complete this sections by discussing two directly verifiable conditions for \mathcal{S}_ρ to be topological.

Lemma 1.3.16

Let ρ be a prametric on Y . and let $y \in \text{int } \mathcal{B}_\varepsilon^\rho(y)$ hold for all $\varepsilon > 0$ and $y \in Y$. Then $\mathcal{S}(\tau_\rho) = \mathcal{S}_\rho$. This is e.g. the case, if the triangle inequality does hold.

Proof: For the first statement it is sufficient to show, that $\rho(y, y_n)$ tends to zero, whenever $y_n \xrightarrow{\tau_R} y$.

Let $\varepsilon > 0$. Since y is an inner point of $\mathcal{B}_\varepsilon^\rho(y)$ there does exist an open set $U \subset Y$, such that $y \in U \subset \mathcal{B}_\varepsilon^\rho(y)$. Since y_n converges to y , it holds that $y_n \in U$ and therefore $\rho(y, y_n) < \varepsilon$ for all n sufficiently large. Therefore $\rho(y, y_n)$ converges to zero.

For the second assertion clearly $\mathcal{B}_\delta^\rho(x) \subset \mathcal{B}_\varepsilon^\rho(y)$ does hold for arbitrary $x \in \mathcal{B}_\varepsilon^\rho(y)$ and $\delta := \frac{1}{2}(\varepsilon - \rho(y, x))$. So the ε -balls are open, and all of its elements are inner points. \square

Remark 1.3.17

Due to Lemma 1.2.15 (ii) and Lemma 1.3.11 unique limits with respect to \mathcal{S}_ρ are also sufficient for $\mathcal{S}(\tau_\rho) = \mathcal{S}_\rho$.

The last remark raises the question, when the sequential convergence structure \mathcal{S}_ρ induced by a prametric actually has unique limits. It is tempting to suspect the implication $\rho(y_1, y_2) = 0 \Rightarrow y_1 = y_2$ to be sufficient for this property. But unfortunately this is wrong, as the subsequent example shows.

Example 1.3.18

Consider $Y = \mathbb{R}^2$ and let $\|\cdot\|$ be an arbitrary norm on \mathbb{R}^2 . Denote by π_i , $i = 1, 2$ the projection on the i -th component, by $\chi_{\mathbb{R} \setminus \{0\}}$ the characteristic function of $\mathbb{R} \setminus \{0\}$, by $\mathbf{1} := (1, 1)^T$ and set $\gamma_i := \chi_{\mathbb{R} \setminus \{0\}} \circ \pi_i$ for $i = 1, 2$.

Then the mapping $\rho : Y \times Y \rightarrow [0, \infty]$ given by

$$\rho(y, \tilde{y}) = \gamma_1(y)|\pi_1(y - \tilde{y})| + \gamma_2(y)|\pi_2(y - \tilde{y})| + \text{eq}(y, \tilde{y}) (\|\tilde{y} - \mathbf{1}\| + 1 - \text{eq}(\mathbf{1}, \tilde{y}))$$

is a prametric on Y , because $|\pi_1(y - \tilde{y})| = |\pi_2(y - \tilde{y})| = \text{eq}(y, \tilde{y}) = 0$ if $\tilde{y} = y$. Now consider $\tilde{y} \neq y$. Then $\text{eq}(y, \tilde{y}) = 1$ and therefore

$$\rho(y, \tilde{y}) = \gamma_1(y)|\pi_1(y - \tilde{y})| + \gamma_2(y)|\pi_2(y - \tilde{y})| + \|\tilde{y} - \mathbf{1}\| + 1 - \text{eq}(\mathbf{1}, \tilde{y}).$$

For $\tilde{y} = \mathbf{1}$ we have $1 - \text{eq}(\mathbf{1}, \tilde{y}) = 1$ and therefore $\rho(y, \tilde{y}) \neq 0$. For $\tilde{y} \neq \mathbf{1}$ we get $\|\tilde{y} - \mathbf{1}\| > 0$ which also yields $\rho(y, \tilde{y}) \neq 0$. Hence $\rho(y, \tilde{y}) = 0$ implies $y = \tilde{y}$.

But for $y = (1, 0)^T$, $\tilde{y} = (0, 1)^T$ and $y_n = (1 + \frac{1}{n}, 1 + \frac{1}{n})^T$ we have

$$\rho(y, y_n) = \rho(\tilde{y}, y_n) = (1 + \|\mathbf{1}\|) \frac{1}{n} \rightarrow 0 \text{ for } n \rightarrow \infty.$$

Due to construction of \mathcal{S}_ρ this yields $y_n \xrightarrow{\mathcal{S}_\rho} y$ and $y_n \xrightarrow{\mathcal{S}_\rho} \tilde{y}$. Because of $y \neq \tilde{y}$, limits with respect to \mathcal{S}_ρ are not unique.

Instead, the property of having unique limits can be characterized as follows.

Lemma 1.3.19

Let ρ be a prametric on Y . Then the following statements are equivalent:

- (i) \mathcal{S}_ρ provides unique limits
- (ii) ρ is separating and for all $y, \tilde{y} \in Y$ and $(y_n) \in \mathcal{S}_\rho(y) \cap \mathcal{S}_\rho(\tilde{y})$ we have $\rho(y, y_n) \rightarrow \rho(y, \tilde{y})$.
- (iii) ρ is separating and $\rho|_{\mathcal{S}_\rho^{-1}\{(y_n)\}^2} \equiv 0$ for all sequences (y_n) in Y such that $\mathcal{S}_\rho^{-1}\{(y_n)\} \neq \emptyset$.

Proof: (i) \Leftrightarrow (ii): First let \mathcal{S}_ρ have unique limits. If there were y and $\tilde{y} \in Y$ such that $y \neq \tilde{y}$ and $\rho(y, \tilde{y}) = 0$, then the constant sequence with the single member \tilde{y} would converge to y and \tilde{y} in contradiction to uniqueness of limits. So ρ is separating. Furthermore, if $(y_n) \in \mathcal{S}_\rho(y) \cap \mathcal{S}_\rho(\tilde{y})$ then uniqueness of limits implies $y = \tilde{y}$ and therefore $\rho(y, y_n) \rightarrow 0 = \rho(y, y) = \rho(y, \tilde{y})$.

For the converse direction assume that (ii) is fulfilled. Let $y, \tilde{y} \in Y$ and $(y_n) \in \mathcal{S}_\rho(y) \cap \mathcal{S}_\rho(\tilde{y})$, i.e. $\lim_{n \rightarrow \infty} \rho(y, y_n) = \lim_{n \rightarrow \infty} \rho(\tilde{y}, y_n) = 0$. Now (ii) implies $\rho(y, y_n) \rightarrow \rho(y, \tilde{y})$ and hence uniqueness of limits in $[0, \infty]$ yields $\rho(y, \tilde{y}) = 0$ and consequently $y = \tilde{y}$. Hence, \mathcal{S}_ρ provides unique limits.

(ii) \Leftrightarrow (iii): Since $\mathcal{S}_\rho^{-1}((y_n)) = \{y \in Y \mid y_n \xrightarrow{\mathcal{S}_\rho} y\}$ the implication (ii) \Rightarrow (iii) has already been shown above. Conversely, consider $(y_n) \in \mathcal{S}_\rho(y) \cap \mathcal{S}_\rho(\tilde{y})$. Then (iii) implies $\rho(y, \tilde{y}) = 0 = \lim_{n \rightarrow \infty} \rho(y, y_n)$ \square

Remark 1.3.20

In [Fle11, Prop. 2.10] it is directly proven, that 1.3.19 (ii) implies topologicality of \mathcal{S}_ρ and that $\tau(\mathcal{S}_\rho)$ is a topology inducing \mathcal{S}_ρ .

Chapter 2

Ill-posed inverse problems and regularization

2.1 Inverse problems and terms of solution

As mentioned in the introduction, inverse problems consist in determining some 'reason' from some measured data, where the relation of the given data to its reason is mathematically modeled by a mapping that maps reasons to data, so retrieving the reason from specific given (exact) data is the same as solving an operator equation.

It will be convenient for the formulation of subsequent terminology, to be able to address all equations given by such a functional relation at once and to regard this entirety as the inverse problem (even if we know, that is often only used for one specific equation). From this point of view, an inverse problem can be formally defined as mapping from the data space to the set of equations given by the operator.

Definition 2.1.1

Let X and Y be sets. The *inverse problem* \mathcal{P}_F associated to an operator $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ is a mapping

$$\mathcal{P}_F : Y \rightarrow \{\text{Equations } Fx = y \mid y \in Y\}$$

given by $y \mapsto Fx = y$ and we will call an equation $\mathcal{P}_F(y)$ an *inverse problem with given right hand side* y .

Having defined, what the inverse problem is, one asks what solving an inverse problem with given right hand side actually means. As it turns out, 'solution' in the strict algebraic sense is often too narrow, be it because the problem is ill-posed, or because 'similarity' in the data space is better described by

a wider notion than strict equality, or by some other reason. Simply re-defining the term solution is often an appropriate mean to cure or at least to alleviate the effects of the purely algebraic points of Hadamard's notion of ill-posedness, i.e. existence and uniqueness. Moreover, an alternative definition of 'solution' may allow for a better model of the real circumstances behind the mathematical representation of the inverse problem or the intended use of such a solution.

In the sequel we will use some additional notation concerning mappings, which we introduce now.

Notation

Let X and Y be sets. Then we denote by $\text{Map}_P(X, Y)$ the set of partial mappings from X to Y , i.e. all mappings $F : \mathcal{D}(F) \subseteq X \rightarrow Y$.

For an extended real valued functional $f : X \rightarrow [0, \infty]$ we denote by $\text{dom}(f) := \{x \in X \mid f(x) < \infty\}$ its effective domain.

We denote the domain of a set-valued mapping $\mathcal{F} : X \rightarrow 2^Y$ by $\text{dom}(\mathcal{F}) := \{x \in X \mid \mathcal{F}(x) \neq \emptyset\}$. By $\text{gr}(\mathcal{F}) := \{(x, y) \in X \times Y \mid x \in X, y \in \mathcal{F}(x)\}$ we denote its graph.

Now, we come to solutions of inverse problems.

Definition 2.1.2

Let X, Y be sets and ρ a prametric on Y .

- (i) For a partial mapping $F \in \text{Map}_P(X, Y)$ and $y \in Y$ we call the set

$$\mathcal{L}_F^{ex, \rho}(y) := \{x \in \mathcal{D}(F) \mid \rho(Fx, y) = 0\}$$

the *set of ρ -solutions* of $\mathcal{P}_F(y)$ and its elements *ρ -solutions* of $\mathcal{P}_F(y)$.

In the special case that $\rho(Fx, y) = 0$ if and only if $Fx = y$ we write

$$\mathcal{L}_F^{ex}(y) := \mathcal{L}_F^{ex, \rho}(y) = \{x \in \mathcal{D}(F) \mid Fx = y\}$$

and call it the *set of exact solutions* of $\mathcal{P}_F(y)$. Its elements are called *exact solutions* of $\mathcal{P}_F(y)$.

- (ii) Let M be a subset of $\text{Map}_P(X, Y)$. We call a mapping

$$\mathcal{L} : M \times Y \rightarrow 2^X, \quad (F, y) \mapsto \mathcal{L}_F(y)$$

term of solution with respect to ρ on M if $\mathcal{L}_F(y) \subseteq \mathcal{L}_F^{ex, \rho}(y)$ for all y with $\mathcal{L}_F^{ex, \rho}(y) \neq \emptyset$.

We then call $\mathcal{L}_F(y)$ *solution set* of $\mathcal{P}_F(y)$ with respect to \mathcal{L} and every element of $\mathcal{L}_F(y)$ is an *\mathcal{L} -solution* of $\mathcal{P}_F(y)$.

If $\mathcal{L}_F^{ex, \rho}(y)$ is the set of exact solutions for all $F \in M$ and $y \in Y$, we will speak of an *exact term of solution*.

-
- (iii) Let \mathcal{L} be a term of solution on $M \subseteq \text{Map}_P(X, Y)$ and $F \in M$. Then we denote by \mathcal{L}_F the mapping $\mathcal{L}_F : Y \rightarrow 2^X$ induced by \mathcal{L} via $y \mapsto \mathcal{L}_F(y) := \mathcal{L}(F, y)$.

Remark 2.1.3

Note, that although we did not assume any additional structure on X and Y in the definitions of an inverse problem and terms of solution, such structure may be used in constructing concrete terms of solutions.

Also, every exact solution is an eq-solution where the prametric eq is defined as in Example 1.3.4.

Remark 2.1.4

Another possibility would be to incorporate the prametric defining the term of solution one intends to use already in the definition of an inverse problems, namely by mapping to the set containing all formal equations of the form $\rho(Fx, y) = 0$. Whether to do so or not, is rather a question of philosophy. One reason to define an inverse problem like in Definition 2.1.1 is the following point of view: On the one hand, there is a physical (financial, ...) process, describing the formation of the observed data out of the quantity of interest more or less accurate, and which is itself mathematically modelled more or less accurate by some mapping, resulting in an (exact) equation. On the other hand, the same physical principle can in general describe a variety of applications, and such an application again can serve a variety of purposes, each having different demands on a 'solution' of the equation given by physics (or whatever). For example, scanning a human being by CT in order to plan a medical intervention differs from scanning a bicycle frame for fissures or from looking into a Kinder egg, just out of curiosity if it contains a Smurf, and that in what aspects of the object one wishes to see as well in the consequences of messy images.

Since choosing an appropriate term of solution includes the decision which solutions are considered 'acceptable' or 'good enough' (or possibly even 'too good') and since this decision will heavily depend on the purpose, it seems reasonable to regard the modelling by an equation and the fixing of the meaning of the word 'solution' as separate parts of the process of designing a method for retrieving the quantity of interest.

Before we remark on sense and nonsense of Definition 2.1.2 we present some examples of terms of solution, which we want to be covered by our definition.

Example 2.1.5

- (i) Let X, Y be arbitrary sets, M an arbitrary subset of $\text{Map}_P(X, Y)$

and ρ a prametric on Y . Then clearly $\mathcal{L}^{ex,\rho} : M \times Y \rightarrow 2^X$ given by $\mathcal{L}^{ex,\rho}(F, y) := \mathcal{L}_F^{ex,\rho}(y)$ is itself a term of solution with respect to ρ .

If $\rho(x, y) = 0$ if and only if $x = y$, then $\mathcal{L}^{ex,\rho} = \mathcal{L}^{ex}$ does hold and consequently every term of solution with respect to ρ is an exact term of solution. In particular, for every metric μ on Y and $p > 0$ the term of solution \mathcal{L}^{ex,μ^p} is an exact term of solution.

- (ii) Let X and Y be two Hilbert spaces and $L(X, Y)$ the set of linear bounded operators from X to Y . Consider

$$\mathcal{L}^\perp : L(X, Y) \times Y \rightarrow 2^X \quad \text{defined by} \quad \mathcal{L}_A^\perp(y) := \mathcal{L}_{A^*A}^{ex}(A^*y),$$

where A^* denotes the Hilbert space adjoint of $A \in L(X, Y)$, i.e. the mapping which maps a pair $(A, y) \in M \times Y$ to the solution set of the associated normal equation or equivalently to $\operatorname{argmin}_{x \in X} \|Ax - y\|_Y$.

We denote by $\pi_{\operatorname{rg}(A)^\perp}$ the projection onto $\operatorname{rg}(A)^\perp$. Then we have

$$\mathcal{L}_A^\perp(y) = \mathcal{L}_A^{ex}(y - \pi_{\operatorname{rg}(A)^\perp}(y)) = \mathcal{L}_A^{ex}(\pi_{\overline{\operatorname{rg}(A)}}(y))$$

which specializes to

$$\mathcal{L}_A^\perp(y) = \begin{cases} \mathcal{L}_A^{ex}(y), & \text{if } y \in \operatorname{rg}(A) \\ \ker(A), & \text{if } y \in \operatorname{rg}(A)^\perp \\ \emptyset, & \text{if } y \in Y \setminus (\operatorname{rg}(A) \oplus \operatorname{rg}(A)^\perp) \end{cases}$$

and hence \mathcal{L}^\perp is an exact term of solution on $L(X, Y)$. For details see e.g. [Rie03] or [EHN96].

- (iii) Let X, Y again be Hilbert spaces. For $A \in L(X, Y)$ we denote by A^\dagger the Moore-Penrose inverse of A . Then the mapping $\mathcal{L}^\dagger : L(X, Y) \times Y \rightarrow 2^X$ given by

$$\mathcal{L}_A^\dagger(y) := \begin{cases} \{A^\dagger(y)\}, & \text{if } y \in \operatorname{rg} A \oplus \operatorname{rg} A^\perp \\ \emptyset, & \text{else} \end{cases}$$

is an exact term of solution. Due to definition (see e.g. [Rie03])

$$A^\dagger(y) = \operatorname{argmin}_{x \in \mathcal{L}_A^\perp(y)} \|x\|_X$$

does hold for $y \in \operatorname{rg}(A) \oplus \operatorname{rg}(A)^\perp$. So, if $y \in \operatorname{rg}(A)$, the unique element of $\mathcal{L}_A^\dagger(y)$ is chosen from $\mathcal{L}_A^{ex}(y)$ and is therefore itself an exact solution.

For arbitrary $x^* \in X$ the same reasoning shows, that $\mathcal{L}^{\dagger, x^*} : L(X, Y) \rightarrow 2^X$ given by

$$\mathcal{L}_A^{\dagger, x^*}(y) := \operatorname{argmin}_{x \in \mathcal{L}_A^{\perp}(y)} \|x - x^*\|$$

is an exact term of solution. So, x^* -minimum norm solutions are also covered by Definition 2.1.2.

- (iv) The previous example can be generalized in the following way. Let X and Y be sets, \mathcal{L} be an arbitrary term of solution with respect to ρ on $M \subseteq \operatorname{Map}_P(X, Y)$ and $R : X \rightarrow \mathbb{R} \cup \{\infty\}$ a functional.

Then the mapping given by

$$(F, y) \mapsto \operatorname{argmin}_{x \in \mathcal{L}_F(y)} R(x)$$

is also a term of solution with respect to ρ . If $\mathcal{L} = \mathcal{L}^{ex, \rho}$ we call the solutions given in this way *R-minimum ρ -solutions* and denote the resulting term of solution by $\mathcal{L}^{\min R, \rho}$.

A popular specimen of this kind of term of solution in a (possibly non-linear) Banach space setting is for fixed $x^* \in X$ given by $\mathcal{L} = \mathcal{L}^{ex}$ and $R(x) = \|\cdot - x^*\|$. We will call it also x^* -minimum-norm solution and denote it by \mathcal{L}^{\min, x^*} .

- (v) In his doctoral thesis [Fle11] J. Flemming chooses a different approach to deal with non-metric similarity measures. There he considers two metric spaces (X, τ_X) , (Y, τ_Y) and an additional third topological space (Z, τ_Z) which he calls data space and contains the possible measurement data, while the operator $F : X \rightarrow Y$ is considered to be sequentially continuous, i.e. F is $\mathcal{S}(\tau_X)$ - $\mathcal{S}(\tau_Y)$ -continuous.

He uses a similarity measure $S : Y \times Z \rightarrow [0, \infty]$ where $\rho(y, z)$ should be small, if the measured data z is a 'good' representation of a right hand side $y \in Y$. Then an element $x \in X$ is called *S-generalized solution* of the equation $\mathcal{P}_F(y)$ if there exists a $z \in Z$, such that $S(Fx, z) = S(y, z) = 0$.

Translated to our notation we get a mapping $\mathcal{L}^{S, Z} : \mathcal{C}_s(X, Y) \times Y \rightarrow 2^X$ given by

$$\mathcal{L}^{S, Z}(F, y) := \mathcal{L}_F^{S, Z}(y) = \{x \in X \mid \exists z \in Z : S(Fx, z) = S(y, z) = 0\}.$$

Now we investigate the question, if $\mathcal{L}^{S, Z}$ is a term of solution in our sense. Unfortunately the related functional $S_Y : Y \times Y \rightarrow [0, \infty]$ given

by

$$S_Y(y_1, y_2) := \inf_{z \in Z} (S(y_1, z) + S(y_2, z))$$

is in general not a prametric, as is pointed out in Example 1.3.4 (iv), and so does not define a term of solution in our sense.

But, clearly $\tilde{S}_Y : Y \times Y \rightarrow [0, \infty]$ given by

$$\tilde{S}_Y(y_1, y_2) := \begin{cases} 0 & \text{if } y_1 = y_2 \\ S_Y(y_1, y_2) & \text{otherwise} \end{cases}$$

is a prametric on Y . Since $x \in \mathcal{L}_F^{S,Z}(y)$ implies $S_Y(Fx, y) = 0$ the inclusion $\mathcal{L}^{S,Z}(F, y) \subseteq \mathcal{L}^{ex, \tilde{S}_Y}$ does hold and hence $\mathcal{L}^{S,Z}$ is a term of solution with respect to \tilde{S}_Y .

If S_Y is itself a prametric on Y then $S_Y = \tilde{S}_Y$ and $\mathcal{L}^{S,Z}$ is a term of solution with respect to S_Y . Under appropriate assumptions on S (see [Fle11, Prop. 2.7]) we even get $\mathcal{L}^{S,Z} = \mathcal{L}^{ex, S_Y}$.

Note that $\mathcal{L}_F^{S,Z}(Fx) = \emptyset$ for $x \in X$ such that $S(Fx, z) \neq 0$ for all $z \in Z$, so without further assumptions on S it is possible that $\mathcal{P}_F(y)$ is solvable in the classical sense, but has no S -generalized solution.

- (vi) Let X and Y be vector spaces over the same field, let M be the set of linear mappings from X to Y and $\mathcal{B} := (B_A)_{A \in M}$ be a family of linear mappings from Y to X such that $AB_A A = A$, i.e. B_A is a generalized or partial inverse of A (see e.g. [BIG03, Nas71]).

Then $\mathcal{L}^{\mathcal{B}} : M \times Y \rightarrow 2^X$ given by $\mathcal{L}^{\mathcal{B}}(A, y) = \{B_A(y)\}$ is an exact term of solution.

Remark 2.1.6

- As to the relation of an arbitrary term of solution to exact solutions, the inclusion $\mathcal{L}_F^{ex}(y) \subseteq \mathcal{L}_F^{ex, \rho}(y)$ does hold for all prametrics ρ on a set Y , $F : X \rightarrow Y$ and $y \in Y$, because $\rho(Fx, Fx) = 0$ for all $x \in X$.

Since this inclusion is allowed to be proper and a term of solution with respect to ρ is just an arbitrary subset of $\mathcal{L}_F^{ex, \rho}(y)$ it can in general not be guaranteed that it even contains an exact solution.

- Our notion allows terms of solution \mathcal{L} , such that $\mathcal{L}_F(y)$ is empty despite $y \in \text{rg } F$. While this behaviour may seem unsatisfying or even undesirable from a mathematical point of view (see e.g. [BIG03]), from a practitioner's perspective it possibly does not matter or even make sense since some right sides may be considered 'impossible' due to physical or other real world reasons.

E.g. in the well-known example of single slice computed tomography modelled by means of the Radon transform, the possible right-hand sides are functions g mapping the pair of real numbers (s, φ) to

$$g(s, \varphi) = \ln \left(\frac{I_0(s, \varphi)}{I(s, \varphi)} \right),$$

where $I_0(s, \varphi)$ is the primary intensity, which would be measured at the detector if not attenuated by an object, and $I(s, \varphi)$ is the intensity attenuated by the actual object along the beam described in parallel beam (or pencil beam) geometry by the distance s to the origin and the rotation angle φ . Unless the object itself emits radiation, a scenario which is usually assumed not to happen, due to the density of the object, the primary intensity will be greater or equal than the attenuated intensity, and so $g \geq 0$. If negative values should occur due to measurement errors, they are usually cut to zero. So, in this case, right sides that take negative values are not of interest. For the physical background see e.g. [Buz08, Chapter 2], a detailed discussion of the mathematical side is e.g. to be found in [Nat86].

- In contrast to strict equality, the relation $\rho(y_1, y_2) = 0$ is not an equivalence relation on Y for general prametrics ρ . Nor does in general the existence of a $y \in Y$ such that $\rho(Fx_1, y) = \rho(Fx_2, y) = 0$ impose an equivalence relation on $\mathcal{D}(F)$.
- We are aware, that it may be desirable to accept elements of X as solutions of an inverse problem, which can maybe not be modelled by prametrics as suggested by us, e.g. the outcome of some approximation method.

2.2 Well-posedness and stability

As already mentioned in the introduction, the concept of well-posedness respectively ill-posedness in the sense of Hadamard makes sense for all inverse problems \mathcal{P}_F defined on sets X and Y equipped with structures which provide an appropriate notion of 'nearness' or getting 'near' and where we can find mappings between them, which respect the respective structure, i.e. there is a sensible notion of 'continuity' between the two structures.

To handle both categories discussed up to now allowing for such notions, namely topological spaces and sequential convergence spaces, simultaneously in up-coming definitions, we subsume them under the term of well-mannered

category. Since we do not want to dwell on category theory but only want to be able to address objects of both categories at the same time, this is done as following.

Definition 2.2.1

We say, that a pair (X, κ) is of *well-mannered category*, if κ is a sequential convergence structure or a topology on X .

Now, we are able to express statements regarding topological continuity and the two defined sorts of continuity involving sequential continuity at once.

Definition 2.2.2

Let (X, κ_X) and (Y, κ_Y) be of well-mannered category. We call a mapping $F : X \rightarrow Y$ $(\kappa_X\text{-}\kappa_Y)$ -continuous, if it is continuous

- (i) in the sense of topological spaces, if κ_X and κ_Y are both topologies,
- (ii) in the sense of sequential convergence spaces, if κ_X and κ_Y are both sequential convergence structures or
- (iii) in the sense of Definition 1.2.5 otherwise.

Remark 2.2.3

Both, Definition 2.2.1 and Definition 2.2.2 are somewhat unsatisfying. Topologies (being subsets of the power set of the underlying set) and sequential convergence structures (being a mapping from the underlying set to the set of sequences in the underlying set) are so different regarding their technical behaviour, that it will be very difficult if not impossible to prove something fundamental going beyond sequences for both categories at once. So, while these definitions allow to formulate statements applying to both categories (or to various mixed settings) simultaneously, every proof of such a statement has to deal with every of the allowed settings separately.

A way to make it better may be to talk of convergence structures instead of topologies (i.e. mappings from the underlying set to the power set of filters on this set, see [BB02]), since the category of topological spaces can be identified with a subcategory of convergence spaces, so we had at least mappings in both cases. Moreover, even if again not every sequential convergence structure comes from a convergence structure, all sequential convergence structures coming from prametrics do (due to [BB02, Prop. 1.7.6]). Following this idea to an satisfying extent (i.e. to more than just additional notation) and adjusting the content of this thesis to that new language (given it works) is left for future work.

Regarding possible ambiguity of the term continuity, we will always state explicitly what is meant if there is the danger of misunderstanding.

Now, having chosen a term of solution one wishes to work with, there may be certain inverse problems, where this step has already done the job, i.e. solutions in the new sense are as benign as we wanted them to be. So, it seems sensible also to incorporate the used term of solution in our concept of well-posedness, as it was for example done to obtain the notion of well-posedness in the sense of Nashed (see Example 2.2.5 below). Doing so and using the previously defined notion of continuity, well-posedness à la Hadamard can be generalized as follows to mappings between objects of well-behaved categories. Due to its global character we will denote it by the term 'globally well-posed'.

Definition 2.2.4

Let (X, κ_X) and (Y, κ_Y) be of well-behaved category, \mathcal{P}_F the inverse problem induced by a mapping $F : X \rightarrow Y$ and \mathcal{L} a term of solution.

- (i) \mathcal{P}_F is called *globally well-posed* with respect to κ_X , κ_Y and \mathcal{L} if
 - (a) The solution set $\mathcal{L}_F(y)$ is non-empty for all $y \in Y$.
 - (b) The cardinality of $\mathcal{L}_F(y)$ is at most one for all $y \in Y$.
 - (c) \mathcal{L}_F considered as single valued mapping is continuous.
- (ii) \mathcal{P}_F is called *globally ill-posed* with respect to κ_X , κ_Y and \mathcal{L} if it is not globally well-posed.

And indeed, in a linear Hilbert space setting and using the Moore-Penrose inverse this specializes to well-posedness in the sense of Nashed.

Example 2.2.5

Let X and Y be Hilbert spaces and $A : X \rightarrow Y$ be a bounded, linear operator. It is known, that the Moore-Penrose inverse $A^\dagger : \text{rg}(A) \oplus \text{rg}(A)^\perp \rightarrow X$ is continuous if and only if $\text{rg}(A)$ is closed in Y (see e.g. [EHN96]). Moreover, A^\dagger is defined on the whole of Y if and only if $\text{rg}(A)$ is closed in Y . Since the term of solution $\mathcal{L}_A^\dagger(y)$ at A and y given by the Moore-Penrose inverse as in Example 2.1.5 is at most single valued anyway, this implies global well-posedness of \mathcal{P}_A with respect to the Hilbert space topologies on X and Y and \mathcal{L}^\dagger if and only if $\text{rg}(A)$ is closed. So, global well-posedness with respect to \mathcal{L}^\dagger is exactly what is commonly known as well-posedness in the sense of Nashed.

There are several different concepts of well-posedness, since well-posedness as defined in Definition 2.2.4 is not completely satisfying. As pointed out e.g. in [Hof00], one of the drawbacks of well-posedness in the sense of Hadamard and with it of Definition 2.2.4 is its purely global character. But often local

properties of a mapping would be sufficient to approximate solutions in an acceptable way. This leads to the following concept of ill-posedness which is frequently used in the context of non-linear inverse problems, e.g by [Hof00] in a Hilbert space setting or by [SKHK12] for the case of Banach spaces.

Example 2.2.6

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces and $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ a mapping. Then the inverse problem \mathcal{P}_F is called *locally ill-posed* at $x_0 \in X$ if for all $r > 0$ there is a sequence (x_n) in $\overline{\mathcal{B}_r(x_0)} \cap \mathcal{D}(F)$ such, that

$$F(x_n) \rightarrow F(x_0) \text{ and } x_n \not\rightarrow x_0.$$

Conversely, \mathcal{P}_F is called *locally well-posed* at $x_0 \in \mathcal{D}(F)$, if it is not locally ill-posed at x_0 . Local well-posedness includes x_0 being an isolated point of the fiber $F^{-1}\{F(x_0)\}$ and the existence of a neighbourhood U of x_0 , such that every selection of $F^{-1}|_{F(U)}$ is continuous in $F(x_0)$ (see [Hof00]).

A detailed discussion of the interdependence of various concepts of ill-posedness in Hilbert spaces plus links to additional literature containing further notions of ill-posedness can be found in [Hof00]. Furthermore, local well-posedness can be generalized to so-called conditional well-posedness, see e.g. [SKHK12, Section 3.1.3] for that.

All of the presented concepts of well-posedness have in common, that they have some minimum requirement on stable approximability of solutions which is after all the real aim we want to achieve.

So, stable approximability, corresponding to the third item of Definition 2.2.4, is the point we will mainly concentrate on. Although it is a desirable property for various reasons, we will not discuss uniqueness (global or local). Instead we will explicitly allow multiple solutions and leave the task of picking the most appropriate solution to constructors of specific settings and approximation procedures.

But dropping the requirement of uniqueness has one serious mathematical drawback: Stability can no longer be simply expressed in terms of continuity as in Definition 2.2.4. So, we have to define, what we understand under the term of stability. Since this term has among other things the function to distinguish problems which need further treatment from problems we consider to be sufficiently manageable, it does not make sense to demand stability properties much stronger than we will accept as outcome of a regularization procedure. Therefore, we decided to fix the term in the following way.

Definition 2.2.7

Let (X, κ_X) and (Y, κ_Y) be of well-mannered category.

-
- (i) A set-valued mapping $\mathcal{G} : Y \rightarrow 2^X$ is called stable with respect to κ_Y and κ_X if for every convergent sequence $(y_n) \in Y$ and $y \in Y$ such that $y_n \xrightarrow{\kappa_Y} y$ and every sequence (x_n) such that $x_n \in \mathcal{G}(y_n)$ the following statements hold:
 - (a) (x_n) converges subsequentially.
 - (b) Every subsequential limit of (x_n) is in $\mathcal{G}(y)$.
 - (ii) Let $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ be a mapping and \mathcal{L} a term of solution. We call the inverse problem \mathcal{P}_F stable with respect to κ_X , κ_Y and \mathcal{L} if $\mathcal{L}_F : Y \rightarrow 2^X$ given by $\mathcal{L}_F(y) = \mathcal{L}(F, y)$ is stable with respect to κ_Y and κ_X .

Remark 2.2.8

Let κ_X be a topology or a sequential convergence structure fulfilling the Urysohn property. If \mathcal{G} is single valued, stability with respect to κ_Y and κ_X is the same as (sequential) κ_Y - κ_X -continuity.

2.3 Regularization

As pointed out in the introduction, there are many important cases, where simply redefining the term of solution does not lead to a sufficiently well-posed new problem. It is, for example, a well-known fact, that the Moore-Penrose inverse of compact bounded linear operator between infinite dimensional Hilbert spaces is always unstable (in the sense of strict continuity).

So, for arriving at acceptable results, we have to take further steps. At this point, the idea of regularization enters the stage. It essentially consists in approximating the given inverse problem by a family of similar, more well-behaved surrogate problems such that the corresponding solution sequence has an appropriate relation to the 'true' solution. Such a family is usually indexed by positive real numbers α , which are also intended to indicate the degree of similarity of original and surrogate problem.

Now, what do 'more well-behaved' and 'appropriate relation to true solutions' actually mean? There are various more or less formal definitions of regularization methods to be found in literature, which in general are closely fitted to very particular settings, i.e. special classes of spaces and mappings inducing the inverse problem. Since they commonly make heavy use of structural properties of these special settings, they can not be readily transferred to more general settings. Therefore it seems reasonable to settle first on a frame of minimum requirements, before discussing some of the mentioned definitions and fixing our definition of regularization afterwards.

Regarding the minimum requirements, we will use part of an informal list of requirements for variational regularization (by which mainly Tikhonov regularization is understood) given in [SGG⁺09, Sec. 3.1] as a guidance, namely three items from that list, which are later on subsumed by the authors under the term 'well-posedness':

- Existence, i.e. every surrogate problem has a solution for all $y \in Y$.
- Stability, i.e. the solutions of every surrogate problem depend in some sense continuously on the data.
- Convergence, i.e. for a sequence of surrogate problems approaching the original problem and a sequence of data approaching exact data, the sequence of corresponding solutions converges in some sense to an exact solution.

Now we see to the promised examples, which are mainly taken from textbooks. In the book [EHN96] a formal definition of regularization is only given for the case of linear operators between Hilbert spaces. Partially translated to our notation from the previous section it reads as follows.

Example 2.3.1 ([EHN96, Def. 3.1])

Let X and Y be Hilbert spaces, $A : X \rightarrow Y$ be linear and bounded, $\alpha_0 \in]0, \infty]$ and let $\mathfrak{A} = (\mathcal{A}_\alpha)_{\alpha \in]0, \alpha_0[}$ be a family of continuous mappings $\mathcal{A}_\alpha : Y \rightarrow X$.

- A mapping $\gamma : \mathbb{R}^+ \times Y \rightarrow]0, \alpha_0[$ is called *parameter choice rule*.
- The family \mathfrak{A} is called *regularization operator* for \mathcal{L}_A^\dagger if for every $y \in \text{dom } \mathcal{L}_A^\dagger$ there exists a parameter choice γ such that

$$\sup\{\|A^\dagger y - \mathcal{A}_{\gamma(\delta, y^\delta)} y^\delta\| \mid y^\delta \in Y, \|y - y^\delta\| \leq \delta\} \rightarrow 0 \text{ for } \delta \rightarrow 0 \quad (2.1)$$

and

$$\sup\{\gamma(\delta, y^\delta) \mid y^\delta \in Y, \|y - y^\delta\| \leq \delta\} \rightarrow 0 \text{ for } \delta \rightarrow 0 \quad (2.2)$$

are fulfilled.

- Let γ be a parameter choice rule and $y \in Y$. Then the pair (\mathfrak{A}, γ) is called *regularization method* for $\mathcal{P}_A(y)$, if (2.1) and (2.2) are fulfilled for this specific y .

While the definition of regularization methods given by Rieder in [Rie03, Def. 3.1.1] is in essence a special case of the definition presented in Example 2.3.1, he provides a generalization to non-linear problems, consisting essentially in switching to a different term of solution.

Example 2.3.2 ([Rie03, Def. 7.3.11])

Let X and Y be Hilbert spaces, $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ continuous. For $x^* \in X$ consider the solution term \mathcal{L}^{\min, x^*} , i.e. x^* -minimum-norm solutions. Further let $\mathfrak{A} = (\mathcal{A}_\alpha)_\alpha > 0$ be a family of continuous mappings $\mathfrak{A}_\alpha : X \times Y \rightarrow X$ and γ a parameter choice with codomain $]0, \infty[$.

Then the pair (\mathfrak{A}, γ) is called *regularization method* for F , if there exists a $x^\dagger \in \mathcal{L}^{\min, x^*}(y)$ such that

$$\sup\{\|x^\dagger - \mathcal{A}_{\gamma(\delta, y^\delta)}(x^*, y^\delta)\| \mid y^\delta \in Y, \|y - y^\delta\| \leq \delta\} \rightarrow 0 \text{ for } \delta \rightarrow 0, \quad (2.3)$$

and (2.2) are fulfilled.

In both examples, stability appears in the form of strict continuity of the components of the regularization operator, while convergence is described by (2.1) and (2.3) respectively. Existence corresponds simply to well-definedness of the mappings \mathcal{A}_α .

In [SKHK12] a definition for a general Banach space setting is given, which is a step towards a unified concept for linear and non-linear problems and leaves as much freedom as possible to adapt the definition to specific situations. Albeit pointing out the importance of stability in some sense for obtaining a usable approximation method in the same chapter, the authors focus mainly on the relation of the approximated solutions to exact solutions. This definition is given in essence literally in the following example.

Example 2.3.3 ([SKHK12, Def. 3.20])

Let X, Y Banach spaces. A mapping that transforms every pair $(y^\delta, \alpha) \in Y \times]0, \bar{\alpha}[$ with $0 < \bar{\alpha} \leq +\infty$ to some well-defined element $x_\alpha^\delta \in X$ is called a regularization (procedure) for the linear operator equation $Ax = y$ [where $x \in X, y \in \text{rg}(A) \subseteq Y$], if there exists an appropriate choice $\alpha = \alpha(y^\delta, \delta)$ of the regularization parameter such that, for every sequence $\{y_n\}_{n=1}^\infty$ with $\|y_n - y\| \leq \delta_n$ and $\delta_n \rightarrow 0$ as $n \rightarrow \infty$, the corresponding regularized solutions $x_{\alpha(y_n, \delta_n)}^{\delta_n}$ converge in a well-defined sense to the solution x^\dagger of $Ax = y$. If the solution is not unique, regularized solutions have to converge to solutions of $Ax = y$ possessing the desired properties, e.g., to minimum norm solutions. For non-linear operator equations $Fx = y$ [where $F : \mathcal{D}(F) \subset X \rightarrow Y, x \in \mathcal{D}(F) \subseteq X, y \in F(\mathcal{D}(F)) \subseteq Y$] with regularized solutions $x_\alpha^\delta \in \mathcal{D}(F)$, the definition is analogous. If the solution of $Fx = y$ is not unique, convergence to solutions possessing desired properties, e.g., to \bar{x} -minimum norm solutions, is required. In case of non-uniqueness, different subsequences of regularized solutions can converge to different solutions of the operator equation, which all possess the same desired property.

As in the examples above, it is desired to incorporate structural behaviour of the operator as linearity or desired properties of solutions into the definition of a regularization procedure, which is achieved by allowing varying terms of solution.

Now, in giving a definition for inverse problems between well-behaved categories, we do not necessarily have access to a noise level δ or some appropriate substitute. Due to the Bakushinskiĭ veto (see [Bak84]), in a classical Hilbert space setting, a parameter choice not depending on the noise level will not lead to a regularization method in the sense of Example 2.3.1. So we have to find a workaround, which allows to incorporate a noise level or something similar if available. So we decided only to demand the existence of sequences of parameters which are supposed to work for one $y \in Y$ and one sequence (y_n) at a time and moreover not to define the term parameter choice rule formally.

Another problem is, that we want to settle on a specific and testable formulation of 'convergence in a well-defined sense'. Choosing convergence in sequential convergence spaces as notion of convergence is too strict, since it would exclude subsequential convergence, which is already widely used. So, loosely following [Sei81] and common practice in variational regularization, we decided on subsequential convergence. Now, our definition of regularization goes as follows.

Definition 2.3.4

Let (X, κ_X) and (Y, κ_Y) be of well-mannered category, let \mathcal{P}_F be the inverse problem given by a mapping $F : \mathcal{D}(F) \subseteq X \rightarrow Y$ and \mathcal{L} a term of solution with respect to a prametric ρ on Y .

We call a mapping

$$\begin{aligned} \mathcal{A} : Y \times]0, \infty[&\rightarrow 2^X \\ (y, \alpha) &\mapsto \mathcal{A}_\alpha(y) \end{aligned}$$

regularization operator regarding \mathcal{L}_F (with respect to κ_X and κ_Y) if the following axioms are fulfilled:

- (R1) Existence: For all $\alpha > 0$ and all $y \in Y$ the set $\mathcal{A}_\alpha(y)$ is non-empty.
- (R2) Stability: For fixed $\alpha > 0$ the mapping $\mathcal{A}_\alpha : Y \rightarrow 2^X$ is stable with respect to κ_Y and κ_X .
- (R3) Convergence: For every $y \in Y$ such that $\mathcal{L}_F(y) \neq \emptyset$ and every sequence $y_n \xrightarrow{\kappa_Y} y$, there exists a sequence $(\alpha_n)_n$ of positive real numbers such that every sequence (x_n) with $x_n \in \mathcal{A}_{\alpha_n}(y_n)$ converges subsequentially with respect to κ_x and every such subsequential limit of (x_n) is in $\mathcal{L}_F(y)$.

Let $y \in Y$. Then a regularization operator \mathcal{A} together with a specific rule to select sequences like in (R3) is called a *regularization method* for $\mathcal{P}_F(y)$ regarding \mathcal{L}_F .

Remark 2.3.5

Let κ_X have the Urysohn-property (i.e. it is a topology or a sequential convergence structure possessing the Urysohn-property) and $y \in Y$ as in (R3). If $\mathcal{L}_F(y)$ is single valued, we get κ_X -convergence to the unique element of $\mathcal{L}_F(y)$ instead of subsequential convergence in (R3).

Remark 2.3.6

When sketching the idea of regularization and stating the convergence requirement in our informal list at the start of this section, we talked about approaching the original problem (i.e. $\mathcal{P}_F(y)$) by surrogate problems (i.e. determining $\mathcal{A}_\alpha(y)$). As in Example 2.3.3 we leave the question open, what this 'approaching' is actually supposed to mean. Approximating the problem is usually described by demanding convergence of the parameter sequence to zero and identifying the original problem with $\alpha = 0$, which, from a conceptual point of view, makes immediately sense if we consider for example the special case of Tikhonov regularization, which will be discussed in the next chapter. However, from a pragmatic point of view, there are cases, where convergence of the parameter sequence is needless since we get convergence without it by assumptions we want to hold anyway for other reasons. For example, in a setting called *exact penalization model* by Burger and Osher, one obtains solutions in the desired sense already for parameters small enough, given a certain source condition (intended to obtain convergence rates) is satisfied (see [BO04] and [HKPS07] for a slightly more general setting).

Example 2.3.7

Every regularization operator in the sense of Example 2.3.1 is a regularization operator regarding \mathcal{L}_A^\dagger . As already mentioned, existence is given as soon as the mappings \mathcal{A} are well-defined. If $y_n \rightarrow y$ in Y , continuity of fixed \mathcal{A}_α implies convergence of $\mathcal{A}_\alpha(y_n)$ to $\mathcal{A}_\alpha(y)$ and consequently convergence of every subsequence to $\mathcal{A}_\alpha(y)$. Finally, for $y_n \rightarrow y \in \text{dom } \mathcal{L}_A^\dagger$ we set $\delta_n := \|y - y_n\|_Y$ and $\alpha_n := \gamma(\delta_n, y_n)$. Then (2.1) implies $\mathcal{A}_{\alpha_n}(y_n) \rightarrow y^\dagger$ for $n \rightarrow \infty$, hence convergence does hold.

Chapter 3

Generalized Tikhonov regularization

3.1 Basic notation

One popular class of regularization methods is given by the concept of generalized Tikhonov regularization, where the regularization operator is obtained by minimization of so called Tikhonov functionals, aiming at R -minimal ρ -solutions. Before speaking of regularization itself in Section 3.2, we introduce some basic notation needed for this kind of regularization methods.

Definition 3.1.1

Let X and Y be sets, $F : \mathcal{D}(F) \subseteq X \rightarrow Y$, $\rho : Y \times Y \rightarrow [0, \infty]$ and $R : X \rightarrow [0, \infty]$

- (i) Let $y \in Y$ and $\alpha > 0$. Then a functional $T_{\alpha,y} : X \rightarrow [0, \infty]$ of the form

$$T_{\alpha,y}(x) := \begin{cases} \rho(Fx, y) + \alpha R(x), & \text{if } x \in \mathcal{D}(F) \\ \infty, & \text{otherwise} \end{cases}$$

is called a Tikhonov functional.

- (ii) In this context ρ is called *discrepancy functional* and R *regularization functional*.
- (iii) The mapping $\mathcal{T} : Y \times]0, \infty[\rightarrow 2^X$

$$\mathcal{T}(y, \alpha) := \mathcal{T}_\alpha(y) := \underset{x \in \mathcal{D}(F)}{\operatorname{argmin}} T_{\alpha,y}(x)$$

is called *Tikhonov operator* (with discrepancy ρ and regularization functional R and given by F) .

Remark 3.1.2

The assumptions that R is bounded below by 0 is superfluous if one has other properties, that guarantee boundedness of R from below as for example sequential compactness of the sublevelsets of R , which are often demanded anyway for other reasons, see e.g. [Fle11, Sec. 2.1].

Here, the discrepancy functional ρ is supposed to measure nearness in the image space and the regularization functional R is meant to stabilize the problem by imposing additional constraints on the solutions of the inverse problem under discussion.

Ideally, the regularization functional should be chosen according to some prior knowledge (or wishes) on the solutions of the actual inverse problem, as sparsity (see [DDD04]) or some specific prior distribution in statistical inversion theory (see e.g. [KS05, Chapters 3 and 5]). Since choice and behaviour of specific regularization terms are not a topic of this thesis, we leave it to a (very incomplete) list of literature dealing with that issue. A small assortment of such treatises consists, besides the two already mentioned texts, e.g. in [Tro06], [Tre10], [BL08] (sparsity), [RA07] (Kullback-Leibler), [ROF92], [RO94], [AV94] (total variation).

Regarding the discrepancy functional, powers of metrics are classical. It is primarily supposed to measure nearness in the data space, but also reflects other properties of the data. For example, MAP estimation provides a direct relation between the nature of noise, as well in relation to its statistical behaviour as to the way it is applied, to certain kinds regularization terms, see again [KS05] for the general background and e.g. [Fle11, Chap. 7], [Fle10], [BB11] and [Pö08] for some specific models. So, the squared L^2 -norm as discrepancy functional corresponds to additively applied Gaussian noise, while the Kullback-Leibler divergence (see Example 5.1.3) is related to Poisson noise.

On the other hand, as already discussed extensively in the introduction and Section 1.3, there are also non-metric functionals, which may be used to describe some sensible kind of nearness. For these reasons, in the last few years also the study of non-metric discrepancy functional shifted into focus.

3.2 Formalization of Tikhonov setups

Since we are aiming at analysing the interplay of the various mathematical objects involved in Tikhonov regularization, it will be convenient to be able to address all the components we are interested in at once. Since the same objects are of interest for other types of variational regularization (i.e. building a regularization operator by using minimization problems as surro-

gate problems), as Ivanov regularization and Morozov regularization (see e.g. [LW13] for these terms), we will subsume them under the term of variational setup, while the term regularizing Tikhonov setup will refer to such a setup from which a Tikhonov regularization method can be built.

Since we are only up to study a very special setting regarding the structures on the involved sets later on (especially in Chapter 4), we confine ourselves to this specific setting from now on, even if the definitions and results in the rest of this chapter could be carried over easily to arbitrary well-behaved categories.

Definition 3.2.1 (Variational setup)

A *variational setup* is a tuple $\mathcal{M} = ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$, consisting of

- sets X and Y , which will serve as underlying sets of the solution space and the data space respectively,
- a topology τ_X on X specifying nearness and convergence in the solution space,
- a sequential convergence structure \mathcal{S} on Y supposed to model vanishing of noise in Y ,
- a topology τ_Y on Y to be used as a technical aid in proofs using techniques from calculus of variations and related to mapping properties of the operator defining the inverse problem in question,
- an inverse problem \mathcal{P}_F given by a sequentially τ_X - τ_Y -continuous mapping $F : \mathcal{D}(F) \subseteq X \rightarrow Y$,
- a parametric ρ on Y measuring similarity in Y , which we call the *discrepancy functional* of \mathcal{M} and
- a functional $R : X \rightarrow [0, \infty]$ called the *regularization functional* of \mathcal{M} , which should model some prior knowledge on the solutions of \mathcal{P}_F , i.e. it gets small if the argument fits well to this prior knowledge.

Remark 3.2.2

We are aware, that in dealing with a concrete inverse problem from practice, the various components of a variational setup are not coequal, since some of them are already given by the model of the underlying real world problem as the sets and the operator, while others as the regularization functional can be chosen in the procedure of regularization itself. As we are only interested in the question, what combinations could work in principle from a purely

mathematical point of view and since this practical distinction is irrelevant for that question, we list them even-handedly.

Definition 3.2.3 (Regularizing Tikhonov setup)

A variational setup $\mathcal{M} = ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ is called a *regularizing Tikhonov setup*, if the Tikhonov operator $\mathcal{T} : Y \times]0, \infty[\rightarrow 2^X$ given by F , ρ and R is a regularization operator regarding $\mathcal{L}_F^{\min R, \rho}$ with respect to τ_X and \mathcal{S} , i.e. \mathcal{T} fulfils the axioms (R1)-(R3) from Definition 2.3.4 which read in this special case as follows:

(R1) Existence: $\operatorname{argmin}_{x \in \mathcal{D}(F)} T_{\alpha, y}(x) \neq \emptyset$ for all $(y, \alpha) \in Y \times]0, \infty[$.

(R2) Stability:

For $\alpha > 0$ fixed, $y_n \xrightarrow{\mathcal{S}} y$ and $x_n \in \operatorname{argmin}_{x \in \mathcal{D}(F)} T_{\alpha, y_n}(x)$, the sequence (x_n) converges subsequentially in τ_X and every subsequential limit of (x_n) is in $\operatorname{argmin}_{x \in X} T_{\alpha, y}(x)$.

(R3) Convergence:

Let $y \in Y$ such that $\mathcal{P}_F(y)$ has an R -minimum ρ -solution and let $y_n \xrightarrow{\mathcal{S}} y$. Then there exists a sequence $(\alpha_n)_{n \in \mathbb{N}}$ of positive real numbers such that every sequence (x_n) satisfying $x_n \in \operatorname{argmin}_{x \in \mathcal{D}(F)} T_{\alpha_n, y_n}(x)$ converges subsequentially in τ_X and every subsequential limit of (x_n) is an R -minimum ρ -solution of $\mathcal{P}_F(y)$.

To make the role of the various components more clear we present the well known example of inpainting via Tikhonov regularization with total variation as regularization functional, taken in this case from [SGG⁺09].

Example 3.2.4 (BV inpainting)

Let $\Omega \subset \mathbb{R}^2$ be a Lipschitz domain (i.e. bounded, open, connected with Lipschitz boundary) and $\Omega_I \subset \mathbb{R}^2$ such that $\overline{\Omega}_I \subset \Omega$. By $\operatorname{BV}(\Omega)$ we denote the set of functions of bounded variation on Ω and by $\operatorname{TV}(u)$ we denote the total variation of u . Now, we aim at Tikhonov regularization in the following setting:

- $X := L^2(\Omega)$ and $Y := L^2(\Omega \setminus \Omega_I)$,
- $F : X \rightarrow Y$ is given by $Fu := u|_{(\Omega \setminus \Omega_I)}$,
- τ_X and τ_Y are the weak topologies on X and Y ,
- \mathcal{S} is the sequential convergence structure of the norm-topology on Y ,
- $\rho : Y \times Y \rightarrow [0, \infty]$ is given by $\rho(z, y) := \|z - y\|_2^2$ and

-
- $R : X \rightarrow [0, \infty]$ is given by

$$R(u) := \begin{cases} \text{TV}(u) & \text{if } u \in \text{BV}(\Omega) \\ \infty & \text{else} \end{cases}.$$

Since F is linear and bounded, it is also weakly continuous, and hence $\mathcal{M} = ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ is a variational setup, and as proven e.g. in [SGG⁺09, Thm. 3.76], \mathcal{M} is even a regularizing Tikhonov setup.

3.3 Some necessary conditions on regularizing Tikhonov setups

In this section we present some necessary conditions on a variational setup to be a regularizing Tikhonov setup, namely on the conditions (R1) and (R2). As to be expected, keeping the components of a variational setting as general as in Section 3.2 we end up with findings of rather trivial character. For considerably stronger results one had to impose additional restrictions on the involved objects, which could for example allow to use characterizations of minimizers as given in [Pö08, Chapter 4].

The first statement deals with the existence axiom, consisting essentially in the observation, that a Tikhonov functional cannot have a minimizer if it is not proper.

Remark 3.3.1 (see also [LW13])

Let $((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ be a variational setup. Then obviously (R1) is fulfilled if and only if $\text{dom } \mathcal{T} = Y \times]0, \infty[$, which implies that $\text{dom } R \cap F^{-1}(\text{dom } \rho(\cdot, y)) \neq \emptyset$ for all $y \in Y$. This again yields the statements $\text{dom } R \cap \mathcal{D}(F) \neq \emptyset$ and $\text{rg } F \cap \text{dom } \rho(\cdot, y) \neq \emptyset$ does hold for all $y \in Y$.

The second result is concerned with topological properties of the sets $\mathcal{T}_\alpha(y)$ and mapping properties of \mathcal{T}_α implied by stability.

Theorem 3.3.2 (see also [LW13])

Let $((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ be a variational setup that fulfils (R2), $\alpha > 0$ and $y \in Y$. Then

- (i) $\mathcal{T}_\alpha(y)$ is sequentially compact and so is $(\bigcup_{n \in \mathbb{N}} \mathcal{T}_\alpha(y_n)) \cup \mathcal{T}_\alpha(y)$ for every sequence (y_n) in Y such that $y_n \xrightarrow{\mathcal{S}} y$.

(ii) The implication

$$\left. \begin{array}{l} y_n \xrightarrow{\mathcal{S}} y \\ x_n \xrightarrow{\tau_X} x \\ x_n \in \mathcal{T}(y_n, \alpha) \end{array} \right\} \Rightarrow x \in \mathcal{T}(y, \alpha)$$

does hold, i.e. the mapping \mathcal{T}_α has sequentially closed graph.

If \mathcal{S} is induced by a topology τ and $\tau \times \tau_X$ is sequential, then $\text{gr}(\mathcal{T}_\alpha)$ is closed for every $\alpha > 0$.

If furthermore \mathcal{T}_α is single valued, then (R2) does hold if and only if \mathcal{T}_α is continuous with respect to \mathcal{S} and the sequential convergence structure of τ_X .

Proof:

- (i) Let (x_n) be a sequence in $\mathcal{T}_\alpha(y)$ and let (y_N) be the constant sequence given by $y_n := y$. Then $y_n \xrightarrow{\mathcal{S}} y$ and $x_n \in \mathcal{T}_\alpha(y_n)$ do hold. Therefore (R2) implies the existence of a convergent subsequence of (x_n) converging to an element of $\mathcal{T}_\alpha(y)$.

To prove the second assertion, consider a sequence $(x_k)_{k \in \mathbb{N}}$ in the union $(\bigcup_{n \in \mathbb{N}} \mathcal{T}_\alpha(y_n)) \cup \mathcal{T}_\alpha(y)$. We distinguish two cases:

- (a) There exists a $\tilde{y} \in \{y_n \mid n \in \mathbb{N}\} \cup \{y\}$ such that $x_k \in \mathcal{T}_\alpha(\tilde{y})$ for infinitely many $k \in \mathbb{N}$. Then (x_k) has a subsequence in $\mathcal{T}_\alpha(\tilde{y})$ and the assertion is covered by the first part of the proof.
- (b) For every $\tilde{y} \in \{y_n \mid n \in \mathbb{N}\} \cup \{y\}$ there are at most finitely many $k \in \mathbb{N}$ such that $x_k \in \mathcal{T}_\alpha(\tilde{y})$. Without loss of generality we can assume that $x_k \notin \mathcal{T}(y)$ for all $k \in \mathbb{N}$, that the x_k are pairwise distinct and that there is at most one $x_k \in \mathcal{T}(y_n)$ for all $n \in \mathbb{N}$. (otherwise we could choose an appropriate subsequence).

Then, the sequence given by $\tilde{y}_k := y_n$ if $x_k \in \mathcal{T}(y_n)$ is well defined and $\{\tilde{y}_k \mid k \in \mathbb{N}\}$ is an infinite subset of $\{y_n \mid n \in \mathbb{N}\}$. Hence (y_n) and (\tilde{y}_k) have a subsequence (\tilde{y}_{k_m}) in common.

Now (\tilde{y}_{k_m}) being a subsequence of (y_n) implies $\tilde{y}_{k_m} \xrightarrow{\mathcal{S}} y$ and due to construction $x_{k_m} \in \mathcal{T}_\alpha(\tilde{y}_{k_m}) = \text{argmin}_{x \in \mathcal{D}(F)} T_{\alpha, \tilde{y}_{k_m}}(x)$ does hold. Applying (R2) yields the existence of a convergent subsequence of (x_{k_m}) with limit in $\mathcal{T}_\alpha(y)$, which completes the proof.

- (ii) The first assertion is just a reformulation of (R2). As to the second assertion, in the case of topological convergence the sequential closedness of \mathcal{T}_α is equivalent to sequential closedness of $\text{gr}(\mathcal{T}_\alpha)$ with respect

to $\tau \times \tau_X$. Since the latter topology is sequential, the assertion follows. Finally, the last assertion is clear since, as a topology, τ_X has the Urysohn-property.

□

Remark 3.3.3

Since we did neither use the τ_X - τ_Y -continuity of F nor the fact, that $\mathcal{T}_\alpha(y)$ is the set of minimizers of a Tikhonov functional (not even that it is a set of minimizers at all), statements and proof of Theorem 3.3.2 apply literally to arbitrary regularization operators.

Finally, there is another observation on the domain of definition of F we want to emphasize for later use.

Remark 3.3.4

Let $\mathcal{M} := ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ be a variational setup. Since $\mathcal{T}_\alpha(y) \subseteq \mathcal{D}(F)$, it is necessary for \mathcal{M} to be a regularizing Tikhonov setup that $\mathcal{D}(F)$ is closed under (subsequential) limits of minimizing sequences (x_n) as in (R2) and (R3).

Chapter 4

Necessary conditions on a common set of sufficient conditions

4.1 A common set of sufficient conditions

In this section we present a common set of sufficient conditions for a variational setup to be a regularizing Tikhonov setup as a motivating example for our subsequent investigations, namely in the form of Theorem 4.1.2. It is used in [Pö08, Chapter 1] and Flemming's study of the latter as a special case of his own (in essence more general) setting in [Fle11, Sec. 2.2] and we modified it by substituting one of the involved topologies by a sequential convergence structure. We will also give a (standard) proof of the regularization properties of this setup, essentially condensed from [Pö08], [HKPS07] and [Fle11], to point out where the various assumptions, which will be the object of further investigation later on, are of importance.

Due to Remark 3.3.4 we will need some kind of restricted closedness condition on $\mathcal{D}(F)$ in any case and moreover, proving existence of minimizers by using the direct method from calculus of variations as proposed in the proof of Theorem 4.1.2 requires closedness of $\mathcal{D}(F)$ under limits of sequences we have even less information about. So, if we allowed $\mathcal{D}(F) \subsetneq X$, imposing some additional restriction to $\mathcal{D}(F)$ would be indispensable. This is usually done by demanding $\mathcal{D}(F)$ to be (sequentially) τ_X -closed (e.g. in [Pö08]), which has the additional advantage of ensuring that all the essential properties of the objects regarding τ_X are passed on to the subspace topology on $\mathcal{D}(F)$ (see [Fle11, Prop. 2.9]). Furthermore, we do not rely on any additional structure by a superset of $\mathcal{D}(F)$ as e.g. of a topological vector space.

So, from this chapter on, we make the following general assumption.

General assumption 4.1.1

From now on, all considered inverse problems are induced by mapping $F : X \rightarrow Y$ fulfilling $\mathcal{D}(F) = X$.

Theorem 4.1.2 (See also [LW13])

Let $\mathcal{M} = ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ be a variational setup fulfilling the following list of assumptions:

- (A1) The sublevelsets $\{x \in X \mid R(x) \leq M\}$ are sequentially compact with respect to τ_X for all $M > 0$, so in particular R is sequentially lower semicontinuous with respect to τ_X
- (A2) $\text{dom } T_{\alpha, y} \neq \emptyset$ for all $y \in Y$ and $\alpha > 0$
- (A3) ρ is sequentially $\tau_Y \times \tau_Y$ lower semicontinuous
- (A4) The sequential convergence structure \mathcal{S} satisfies the conditions

$$\mathcal{S} = \mathcal{S}_\rho, \text{ i.e. } y_n \xrightarrow{\mathcal{S}} y \text{ if and only if } \rho(y, y_n) \rightarrow 0 \quad [\text{CONV}]$$
 and

$$\rho(y, y_n) \rightarrow 0 \text{ implies } \rho(z, y_n) \rightarrow \rho(z, y) \text{ for all } z \in \text{dom } \rho(\cdot, y)$$
- (A5) $y_n \xrightarrow{\mathcal{S}} y$ implies $y_n \xrightarrow{\tau_Y} y$

Then \mathcal{M} is a regularizing Tikhonov setup.

Before proving Theorem 4.1.2, we formulate a little observation as lemma. Though being rather trivial in proof itself, it is the key argument in the proof of Theorem 4.1.2 and will be subject to deeper study in the further course of this thesis.

Lemma 4.1.3

Let \mathcal{M} be a variational setup as in Theorem 4.1.2. Then the condition

$$y_n \xrightarrow{\mathcal{S}} y \text{ implies } \rho(z, y_n) \rightarrow \rho(z, y) \text{ for all } z \in \text{dom } \rho(\cdot, y) \quad [\text{CONT}]$$

does hold.

Proof of Lemma 4.1.3: Let (y_n) be a sequence in Y such that $y_n \xrightarrow{\mathcal{S}} y$. Due to [CONV] we have $\rho(y, y_n) \rightarrow 0$ and hence the second statement of (A4) yields $\rho(z, y_n) \rightarrow \rho(z, y)$ for all $z \in \text{dom } \rho(\cdot, y)$. \square

Proof of Theorem 4.1.2: We will only give a sketch of the proof, for details we refer to [HKPS07, Pö08, Fle11]. We have to show, that (R1)–(R3) as in Definition 3.2.3 are fulfilled.

Consider arbitrary $(y, \alpha) \in Y \times]0, \infty[$. Since ρ and R are non-negative $T_{\alpha,y}$ is bounded from below. Therefore (A1)–(A3) allow to use the direct method from the calculus of variations for proving $\mathcal{T}_\alpha(y) \neq \emptyset$ for all $(y, \alpha) \in Y \times]0, \infty[$ and hence (R1) does hold.

For (R2) consider $\alpha > 0$, $y_n \xrightarrow{\mathcal{S}} y$ and $x_n \in \mathcal{T}_\alpha(y_n)$. Due to (A2) and [CONT], there exists $x_0 \in X$ such that $(T_{\alpha,y_n}(x_0))_{n \in \mathbb{N}}$ converges to $T_{\alpha,y}(x_0)$ and hence, is bounded. Now, $x_n \in \operatorname{argmin}_{x \in X} T_{\alpha,y_n}(x)$ implies boundedness of $(R(x_n))_{n \in \mathbb{N}}$. Hence, (A1) delivers a convergent subsequence.

Now, let \bar{x} be a limit of such a subsequence (x_{n_k}) and consider arbitrary $x \in \operatorname{dom} \mathcal{T}_\alpha(y)$. Using (A5), (A3), (A1) and the definition of the x_n we get $T_{\alpha,y}(\bar{x}) \leq \liminf_{k \rightarrow \infty} T_{\alpha,y_{n_k}}(x)$. Thus, [CONT] yields $\bar{x} \in \mathcal{T}_\alpha(y)$ and \mathcal{T}_α is stable.

For (R3) let $x^\dagger \in \mathcal{L}^{\min R, \rho}(y)$, $y_n \xrightarrow{\mathcal{S}} y$ and $x_n \in \mathcal{T}_\alpha(y_n)$. Set

$$\alpha_n := \begin{cases} \sqrt{\rho(Fx^\dagger, y_n)}, & \text{if } \rho(Fx^\dagger, y_n) \in]0, \infty[\\ \frac{1}{n}, & \text{otherwise} \end{cases}. \quad (4.1)$$

Due to $x^\dagger \in \mathcal{L}^{\min R, \rho}(y)$ we have $\rho(Fx^\dagger, y) = 0$ and hence, [CONT] implies $\rho(Fx^\dagger, y_n) \rightarrow 0$. This again implies

$$\alpha_n \rightarrow 0 \text{ and } \frac{\rho(Fx^\dagger, y_n)}{\alpha_n} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (4.2)$$

Now we have $R(x_n) \leq \frac{1}{\alpha_n} T_{\alpha_n, y_n}(x^\dagger)$ for $x_n \in \operatorname{argmin}_{x \in X} T_{\alpha_n, y_n}(x)$. Hence (4.2) together with (A1) yields subsequential convergence of (x_n) and $R(\bar{x}) \leq R(x^\dagger)$ for every subsequential limit of (x_n) .

Furthermore, $\alpha_n \rightarrow 0$ and $\rho(Fx^\dagger, y_n) \rightarrow 0$ imply $\rho(Fx_n, y_n) \rightarrow 0$. Hence, (A3) and (A5) deliver $\rho(F\bar{x}, y) = 0$ and consequently $\bar{x} \in \mathcal{L}^{\min R, \rho}(y)$ for every subsequential limit \bar{x} of (x_n) . \square

Remark 4.1.4

- The parameter choice (4.1) is also discussed in [Fle11, Rem. 3.5].
- The proof of (R3) provides only a statement of existence regarding an appropriate parameter choice, since the parameter choice rule given in (4.1) depends already on an (unknown) solution of \mathcal{P}_F and may therefore be unfeasible for immediate practical use.

If $\mathcal{L}^{\min R, \rho}(y)$ is an exact term of solution (and hence $Fx^\dagger = y$) or ρ obeys a generalized triangle inequality, then Fx^\dagger can be substituted by y in (4.1), still having the drawback of depending on unknown exact data. This can be overcome, if appropriate noise levels are known.

4.2 CONV and CONT

As we have seen in the proofs of Theorem 4.1.2 and Lemma 4.1.3, the properties [CONV] and [CONT] are essential for standard proof techniques to work.

Even if the proof of Theorem 4.1.2 works fine for arbitrary sequential convergence structures \mathcal{S} fulfilling the demanded assumptions, it is common to work in a purely topological framework. Usually the sequential convergence structure \mathcal{S} is assumed to be topological or there are additional assumptions (as ρ being a separating prametric together with some other conditions) assumed which guarantee topologicality in the first place as in [Pö08] and [Fle11, Prop. 2.10].

So, it is an interesting question to ask, when there actually is a topology on Y that satisfies both properties simultaneously. This will be the subject of this section. To make the verbalization of our analysis easier, we make the following definition.

Definition 4.2.1

Let ρ be a sequential convergence structure on Y . A topology τ on Y is said to fulfil

- (i) [CONV] if $\mathcal{S}(\tau)$ fulfils [CONV] as in (A4) and
- (ii) [CONT] if $\mathcal{S}(\tau)$ fulfils [CONT] as in Lemma 4.1.3.

Obviously, asking if there is a topology fulfilling [CONV] is the same as asking if the sequential convergence structure \mathcal{S}_ρ induced by the prametric ρ is topological, a question which has been discussed in length in Section 1.3. As we have seen there, the prametric topology τ_ρ has a singled out position between all topologies whose sequential convergence structures are weaker than \mathcal{S}_ρ , i.e. it is finer than every such topology. In particular, it is maximal with respect to inclusion between all topologies inducing \mathcal{S}_ρ , if there is any at all.

Now the idea is to construct a second topology, which has a similarly distinguished role between all topologies fulfilling [CONT], more precisely a topology that is inclusion minimal with this property. Then, a topology τ fulfilling both properties could be sandwiched between τ_ρ and this second topology, a circumstance which should give some deeper insight on how such a topology τ has to look like.

4.2.1 Bottom slice topologies

In this subsection we will construct the previously promised topology, which, besides being itself a candidate for a topology inducing the sequential convergence structure of a Tikhonov regularization setup, will serve as a tool for analysing topologies fulfilling [CONT].

Taking a closer look on the condition [CONT], one is immediately reminded on sequential continuity of $\rho(z, \cdot)$ in all points $y \in \text{dom } \rho(z, \cdot)$. Therefore, the idea of using some sort of initial topology (see 1.1.8) seems to suggest itself.

Definition 4.2.2 (Bottom slice topology, see also [LW13])

Let $\rho : Y \times Y \rightarrow [0, \infty]$ be a prametric, $Z \subseteq Y$, $\tilde{Y} \subseteq Y$ non-empty and let $[0, \infty]$ be equipped with the order topology.

For $z \in Z$ define

$$f_z : \tilde{Y} \rightarrow [0, \infty] \text{ by } f_z(\tilde{y}) := \rho(z, \tilde{y}).$$

By $\tau_{IN}^{Z, \rho}$ we denote the initial topology on \tilde{Y} with respect to the family $(f_z)_{z \in Z}$ and call it the bottom slice topology with respect to ρ and Z on \tilde{Y} . In the case $\tilde{Y} = Z = Y$ we write $\tau_{IN}^\rho := \tau_{IN}^{Y, \rho}$.

For the sake of ease of notation we make the following general assumption.

General assumption 4.2.3

For the rest of this chapter all subsets of Y denoted by Z or \tilde{Y} will be assumed to be non-empty.

Before we take a closer look on such topologies and their properties, there are some comments on the notation.

Remark 4.2.4

The idea behind using this construction is that of interpreting the property [CONT] as continuity of all of the inducing functionals f_z , $z \in Y$ and deducing that a topology having [CONT] has to be coarser than τ_{IN}^ρ because it is, due to its construction as initial topology, the coarsest topology having [CONT]. Unfortunately, this elementary approach will not work out in general due to various technical restrictions imposed by our setting. They will be more clear after having seen some of the basic properties of bottom slice topologies and will therefore be discussed in detail in Remark 4.2.9. Nevertheless, allowing \tilde{Y} and Z to be proper and distinct subsets of Y , we will be able to derive necessary conditions in terms of bottom slice topologies on subspaces for topologies to induce sequential convergence structures satisfying assumption (A4) (see Section 4.2.2 for this).

Besides the technical reasons for introducing Z and \tilde{Y} separately, they also make sense from a philosophical point of view. As was already mentioned in [LW13], from the practical side in inverse problems, both components of the discrepancy functional ρ have a slightly different interpretation: The first corresponds primarily to exact data and the second to possible measurement data, which (e.g. due to properties of the operator or specifics of the measurement process) may exhibit different characteristics, for example the range of the operator could be a vector space of functions while the measurement data consists only of non-negative functions (cf. the CT example in Remark 2.1.6). This interpretation will have no practical relevance in the restricted framework of this thesis, since we demanded ρ to be a prametric (and has therefore to act on the entire set $Y \times Y$) but we only get a topology on \tilde{Y} . A fruitful way to exploit this point of view, would be to consider data space and measurement space as mostly unrelated and let the discrepancy functional act on their Cartesian product, as is done in the more general setting considered in [Fle11]. It would be an interesting question to relate the therein involved topologies to topologies built in analogy to bottom slice topologies. This subject is left to future work.

Now, back on topic, we state how bottom slice topologies can be constructed explicitly. Afterwards we will give some examples of bottom slice topologies.

Remark 4.2.5

Let Y , ρ , \tilde{Y} and Z be as in Definition 4.2.2. For $z \in Z$ and $a, b \in [0, \infty]$ we have

$$f_z^{-1}([0, b]) = \mathcal{B}_b^\rho(z) \cap \tilde{Y} \text{ and } f_z^{-1}(]a, \infty]) = \{y \in \tilde{Y} \mid \rho(y, z) > a\}.$$

So, due to Lemma 1.1.9 and Lemma 1.1.11, the bottom slice topology is generated by all such sets.

Example 4.2.6

- (i) Let Y be a normed space with norm $\|\cdot\|$ and $p > 0$. Consider $\rho(z, y) = \|z - y\|^p$ and $Z = \tilde{Y} = Y$. Then clearly, the bottom slice topology with respect to ρ and Y is the norm topology.
- (ii) Let $Y := \mathbb{R}^2$ and $Z = \tilde{Y} = Y$. As in Remark 1.3.8 (i), we consider the prametric ρ on Y given by

$$\rho(z, y) = \begin{cases} 0, & \text{if } \#\{i \mid i = 1, 2, y_i = z_i\} \geq 1 \\ 1, & \text{otherwise} \end{cases}$$

for $z = (z_1, z_2)$ and $y = (y_1, y_2)$.

Then the bottom slice topology τ_{IN}^ρ is the initial topology induced by the family $(f_z)_{z \in Y}$ given by $f_z(y) := \rho(z, y)$ for all $y \in Y$. For $z \in Y$ and $a, b \in [0, \infty]$ we have

$$f_z^{-1}([0, b]) = \begin{cases} \mathbb{R}^2 & \text{if } b > 1 \\ f_z^{-1}\{0\} = \mathcal{B}_1^\rho(z) = (\{y_1\} \times \mathbb{R}) \cup (\mathbb{R} \times \{y_2\}) & \text{else} \end{cases}$$

and

$$f_z^{-1}([a, \infty]) = \begin{cases} \emptyset & \text{if } a > 1 \\ f_z^{-1}\{1\} = \mathbb{R}^2 \setminus f_z^{-1}\{0\} & \text{else} \end{cases}$$

So, the set

$$\{\mathcal{B}_1^\rho(z) \mid z \in Y\} \cup \{\mathbb{R}^2 \setminus \mathcal{B}_1^\rho(z) \mid z \in Y\}$$

is a subbase for τ_{IN}^ρ . Now, let $y = (y_1, y_2) \in Y$ be arbitrary. Consider $z^{(1)} = (y_1, y_2 + 1)$, $z^{(2)} := (y_1 + 1, y_2)$ and $z^{(3)} = (y_1 + 1, y_2 + 1)$. Then we get

$$\{y\} = \mathcal{B}_1^\rho(z^{(1)}) \cap \mathcal{B}_1^\rho(z^{(2)}) \cap (\mathbb{R}^2 \setminus \mathcal{B}_1^\rho(z^{(3)}))$$

and hence, all singletons are open. Therefore, τ_{IN}^ρ is the discrete topology on \mathbb{R}^2 .

In particular, τ_{IN}^ρ is different from τ_ρ , which is the trivial topology as has been shown in Remark 1.3.8, and its convergent sequences are exactly the sequences being constant from some index on.

(iii) See Section 5.2 for a class of further examples.

One of the main features we are interested in is sequential convergence in bottom slice topologies. Since they are initial topologies, sequential convergence looks as follows.

Lemma 4.2.7 (see also [LW13])

- (i) For arbitrary subsets Z and \tilde{Y} of Y the sequential convergence structure $\mathcal{S}(\tau_{IN})$ on \tilde{Y} induced by $\tau_{IN}^{Z, \rho}$ can be characterized as follows:

$$y_n \xrightarrow{\mathcal{S}(\tau_{IN}^{Z, \rho})} y \text{ if and only if } \rho(z, y_n) \rightarrow \rho(z, y) \text{ for all } z \in Z.$$

- (ii) If $\tilde{Y} \subseteq Z$ does hold, then $\mathcal{S}(\tau_{IN}^{Z, \rho})$ is stronger than $\mathcal{S}_{\rho|_{\tilde{Y} \times \tilde{Y}}}$.

Proof:

- (i) This is exactly the statement of Lemma 1.1.10 for this special case.
- (ii) Let $y_n, y \in \tilde{Y}$ such that $y_n \xrightarrow{\tau_{IN}^{Z,\rho}} y$. Then item (i) of this lemma implies $\rho(y, y_n) \rightarrow \rho(y, y) = 0$.

□

So, it is quite clear, that we are able to come up to one of our aims, namely to find a candidate for a topology on Y inducing a sequential convergence structure as in (A4), by choosing an appropriate set Z . Properly formulated we arrive at the following corollary.

Corollary 4.2.8

Let ρ be a prametric on Y . Then the topology τ_{IN}^ρ on Y fulfils [CONT].

As already indicated in Remark 4.2.4, Lemma 4.2.7(i) demonstrates that it is not as simple to arrive at the objective of finding a minimal topology with [CONT] by using bottom slice topologies. We discuss this issue in detail now.

Remark 4.2.9

- The condition [CONT] only refers to sequences. Since we need continuity of the f_z for the approach sketched in Remark 4.2.4, the latter will only work for topologies where continuity and sequential continuity coincide, a feature we can only guarantee for sequential topologies.
- As soon as there is one y in Y such that $\text{dom } \rho(\cdot, y) \subsetneq Z$ then even sequential continuity of all the f_z , $z \in Z$ is a stronger demand than [CONT], i.e. $\tau_{IN}^{Z,\rho}$ on Y may be too fine for our purpose.
- As soon as there is one y in Y such that $Z \subsetneq \text{dom } \rho(\cdot, y)$ then the bottom slice topology $\tau_{IN}^{Z,\rho}$ on Y may not even fulfil [CONT] itself, i.e. it may be too coarse for our purpose.
- To sum it up, we can guarantee that the bottom slice topology $\tau_{IN}^{Z,\rho}$ on Y has [CONT] and that $\mathcal{S}(\tau_{IN}^{Z,\rho})$ is weaker than every (topological) sequential convergence structure fulfilling [CONT] at the same time alone by the naive use of the proof idea from Remark 4.2.4 if and only if $Z = \text{dom } \rho(\cdot, y)$ for all $y \in Y$. Unfortunately, the last condition is equivalent to finiteness of ρ , since we have $y \in \text{dom } \rho(\cdot, y)$ for all $y \in Y$, which is a severe restriction on the class of prametrics an analysis based on the global minimality of a bottom slice topology can be applied to.

As already mentioned, variation of \tilde{Y} and Z allows us to derive sensible results by such elementary means despite all that problems, as for example the following lemma.

Lemma 4.2.10

Let ρ be a prametric on Y and let τ be a topology on Y satisfying [CONT]. Further let $\tilde{Y}, Z \subseteq Y$ be such that $\tau|_{\tilde{Y}}$ is sequential and $Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \text{dom } \rho(\cdot, \tilde{y})$. Then $\tau_{IN}^{Z, \rho}$ is coarser than $\tau|_{\tilde{Y}}$.

Proof: Let $z \in Z$ and $\tilde{y}_n, \tilde{y} \in \tilde{Y}$ such that $\tilde{y}_n \xrightarrow{\tau|_{\tilde{Y}}} \tilde{y}$. Due to definition of the subspace topology $\tilde{y}_n \xrightarrow{\tau} \tilde{y}$ does also hold. Using [CONT] and $Z \subseteq \text{dom } \rho(\cdot, \tilde{y})$ we get $\rho(z, \tilde{y}_n) \rightarrow \rho(z, \tilde{y})$. So, all the $f_z := \rho(z, \cdot)|_{\tilde{Y}}, z \in Z$ are sequentially continuous with respect to $\tau|_{\tilde{Y}}$. Since $\tau|_{\tilde{Y}}$ is sequential, they are also continuous and therefore $\tau_{IN}^{Z, \rho}$ is coarser than $\tau|_{\tilde{Y}}$. \square

Remark 4.2.11

Note that in Lemma 4.2.10 the inclusion $\tilde{Y} \subsetneq Z$ is possible because $\text{dom } \rho(\cdot, \tilde{y})$ is the effective domain of $\rho(\cdot, \tilde{y})$ in the entire set Y .

4.2.2 Necessary conditions

Now we are able to deduce some necessary conditions on the existence of topologies satisfying both of the conditions [CONV] and [CONT] in terms of bottom slice topologies and prametric topologies.

Since every topology satisfying [CONV] has a sequential convergence structure which is weaker than the sequential convergence structure induced by τ_ρ , we first attend to the question, what it does mean for τ_ρ if there exists any such topology satisfying additionally [CONT].

Lemma 4.2.12 (See also [LW13])

Let ρ be a prametric on Y and let there be a topology τ on Y fulfilling [CONT] such that $\mathcal{S}(\tau)$ is weaker than $\mathcal{S}(\rho)$. Then τ_ρ also fulfils [CONT].

Proof: Due to Theorem 1.3.13 τ is coarser than τ_ρ . So $y_n \xrightarrow{\tau_\rho} y$ implies $y_n \xrightarrow{\tau} y$, which again yields $\rho(z, y_n) \rightarrow \rho(z, y)$ for all $z \in \text{dom } \rho(\cdot, y)$. \square

In consequence of this lemma, every necessary condition derived from τ_ρ having [CONT] is a necessary condition for the existence of any topology satisfying [CONV] and [CONT] at the same time. So, it is again τ_ρ we put to further investigation.

Theorem 4.2.13

Let ρ be a prametric on Y and suppose that the prametric topology τ_ρ on Y has [CONT].

- (i) For all $\tilde{Y} \subseteq Y$ and for all $Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \text{dom } \rho(\cdot, \tilde{y})$ the bottom slice topology $\tau_{IN}^{Z, \rho}$ on \tilde{Y} is coarser than $\tau_{\rho|(\tilde{Y} \times \tilde{Y})}$.
- (ii) For all open or closed subsets $\tilde{Y} \subseteq Y$ and for all $Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \text{dom } \rho(\cdot, \tilde{y})$ the bottom slice topology $\tau_{IN}^{Z, \rho}$ on \tilde{Y} is coarser than $(\tau_\rho)|_{\tilde{Y}} = \tau_{\rho|(\tilde{Y} \times \tilde{Y})}$.
- (iii) For all $\tilde{Y}, Z \subseteq Y$ such that $\tilde{Y} \subseteq Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \text{dom } \rho(\cdot, \tilde{y})$ the topologies $\tau_{IN}^{Z, \rho}$ and $\tau_{\rho|(\tilde{Y} \times \tilde{Y})}$ on \tilde{Y} have the same convergent sequences. Namely, $\mathcal{S}(\tau_{IN}^{Z, \rho}) = \mathcal{S}(\tau_{\rho|(\tilde{Y} \times \tilde{Y})}) = \mathcal{S}_{\rho|(\tilde{Y} \times \tilde{Y})}$ does hold.

Proof:

- (i) Due to Lemma 1.3.9 the prametric topology $\tau_{\rho|(\tilde{Y} \times \tilde{Y})}$ on \tilde{Y} is finer than the subspace topology $(\tau_\rho)|_{\tilde{Y}}$. Therefore convergence with respect to $\tau_{\rho|(\tilde{Y} \times \tilde{Y})}$ implies convergence with respect to $(\tau_\rho)|_{\tilde{Y}}$. Consequently, [CONT] yields sequential continuity of all $f_z := \rho(z, \cdot)|_{\tilde{Y}}$ and hence continuity, as was to be demonstrated.
- (ii) Let \tilde{Y} be open or closed and $Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \text{dom } \rho(\cdot, \tilde{y})$. Then Lemma 1.3.9 implies $(\tau_\rho)|_{\tilde{Y}} = \tau_{\rho|(\tilde{Y} \times \tilde{Y})}$. Now, 4.2.13(i) (or alternatively sequentiality of $(\tau_\rho)|_{\tilde{Y}}$ combined with Lemma 4.2.10) proves the claim.
- (iii) Let $\tilde{y} \in \tilde{Y}$ and let $(\tilde{y}_n)_{n \in \mathbb{N}}$ be a sequence in \tilde{Y} such that $\rho|(\tilde{Y} \times \tilde{Y})(\tilde{y}, \tilde{y}_n) = \rho(\tilde{y}, \tilde{y}_n) \rightarrow 0$. Due to Lemma 1.3.7 this implies $\tilde{y}_n \xrightarrow{\tau_{\rho|(\tilde{Y} \times \tilde{Y})}} \tilde{y}$ and hence $y_n \xrightarrow{\tau_{IN}^{Z, \rho}} y$. Now, $\tilde{y}_n \xrightarrow{\tau_{IN}^{Z, \rho}} \tilde{y}$ implies $\rho(\tilde{y}, \tilde{y}_n) \rightarrow 0$ due to Lemma 4.2.7(ii).

□

Specialized to the case $\tilde{Y} = Y$ this reads as follows.

Corollary 4.2.14 (See also [LW13, Thm. 3.17])

Let ρ be a prametric on Y . If τ_ρ fulfils [CONT] the following assertions are true:

- (i) For all $Z \subseteq \bigcap_{y \in Y} \text{dom } \rho(\cdot, y)$ the bottom slice topology $\tau_{IN}^{Z, \rho}$ is coarser than τ_ρ .
- (ii) If ρ is finite then τ_ρ and τ_{IN}^ρ both satisfy [CONV]. In particular $\mathcal{S}(\tau_{IN}^\rho) = \mathcal{S}(\tau_\rho) = \mathcal{S}_\rho$ does hold.

Remark 4.2.15

The results previously presented in this section do hold for all Z such that all the $f_z := \rho(z, \cdot)|_{\tilde{Y}}$ are $\tau_{\rho|_{(\tilde{Y} \times \tilde{Y})}}$ -continuous for some other reason too, even, if the assumption $Z \subseteq \bigcap_{\tilde{y} \in \tilde{Y}} \rho(\cdot, \tilde{y})$ is violated.

Finally, we can deduce for a special case a necessary condition on a topology to have [CONV] and [CONT] itself.

Corollary 4.2.16

Let ρ be a finite prametric on Y .

- (i) If there is a topology τ on Y fulfilling [CONV] and [CONT] simultaneously, then $\mathcal{S}(\tau) = \mathcal{S}(\tau_\rho) = \mathcal{S}(\tau_{IN}^\rho)$.
- (ii) τ_ρ is the only sequential topology which can fulfil both conditions at once.

Proof:

- (i) Combine Lemma 4.2.12 and Corollary 4.2.14.
- (ii) If there is a sequential topology fulfilling [CONV], then, according to Theorem 1.3.13 so does τ_ρ . Since both topologies are sequential and have the same convergent sequences, they coincide.

□

The last statement in this section is, in contrast to the previous results, also applicable to non-topological sequential convergence structures and deals with the interdependence of two of the assumptions of Theorem 4.1.2.

Remark 4.2.17 (See also [LW13])

Let $\mathcal{M} = ((X, \tau_X), (Y, \mathcal{S}, \tau_Y), \mathcal{P}_F, \rho, R)$ be a variational setup such that (A4) of Theorem 4.1.2 does hold.

Then (A5) does hold if and only if τ_Y is coarser than τ_ρ .

Proof: Since $\mathcal{S} = \mathcal{S}_\rho$, this is covered by Theorem 1.3.13(i).

□

Chapter 5

The special case of Bregman discrepancies

In this chapter we will apply some of the theory from the previous chapters to prametrics which are built from so called Bregman distances. In the inverse problems context, Bregman distances are mainly used to derive (and express) convergence rates (e.g. [Res05]). Since they exhibit a nice relation to certain kinds of probability distributions (even non-Gaussian ones, see e.g. [Pö08, Sec. 2.2] for details), they seem to be an appropriate tool to describe the amount of noise in a way that respects its statistical nature. So, despite being in general non-metric, their potential role as a discrepancy measure shifted into focus at least in the last decade (see e.g. [RA07] and again [Pö08]).

For defining Bregman distances and related terms and for formulating subsequent statements in this chapter, we will use some additional notation.

Notation

Let V be a real Banach space and let $J : Y \rightarrow \mathbb{R} \cup \{\infty\}$ be a mapping. We denote by

$$\partial J(w) := \{\xi \in V^* \mid J(w) + \langle \xi, v - w \rangle \leq J(v) \forall v \in V\}$$

the subdifferential of J , and by $\langle \cdot, \cdot \rangle$ the dual pairing on $V^* \times V$.

Moreover, we denote by $\text{span}(M)$ the linear subspace of V generated by a subset M of V and for $y_n^*, y^* \in V^*$, we will write $y_n^* \xrightarrow{*} y^*$ for weak* convergence of (y_n^*) to y^* .

5.1 Bregman distances - definition and basic properties

First we define, what we are speaking about in this chapter.

Definition 5.1.1 (Generalized Bregman distance)

Let V be a real Banach space and $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex.

- (i) For $y \in \text{dom } \partial J$ let $y^* \in \partial J(y)$. Then the (*generalized*) *Bregman distance* of y to $z \in V$ with respect to J and y^* is defined as

$$D_J^{y^*}(z, y) := J(z) - J(y) - \langle y^*, z - y \rangle$$

- (ii) Let $Y \subseteq V$ be a subset of V and $\phi : V \rightarrow V^*$ be a selection of ∂J , i.e. $\phi(y) \in \partial J(y)$ for all $y \in Y$. Then we call the mapping $D_J^\phi : Y \times Y \rightarrow [0, \infty]$ given by

$$D_J^\phi(z, y) := \begin{cases} D_J^{\phi(y)}(z, y) & \text{if } y \in \text{dom } \partial J \cap Y \\ 0 & \text{if } y \notin \text{dom } \partial J \text{ and } y = z \\ \infty & \text{else} \end{cases}$$

the ϕ -selected Bregman prametric on Y .

If ∂J is at most single valued on Y we will denote the restriction of any selection ϕ of ∂J to $Y \cap \text{dom } \partial J$ by ∇J . We denote the (then unique) Bregman prametric on Y by D_J^∇ and call it simply *Bregman prametric* on Y . So for $y \in \text{dom } \partial J$ we have

$$D_J^\nabla(z, y) = J(z) - J(y) - \langle \nabla J(y), z - y \rangle.$$

- (iii) Let $Y \subseteq V$ be a subset of V and $\phi : V \rightarrow V^*$ be a selection of ∂J . Then we call the functional ${}_J^\phi D : Y \times Y \rightarrow [0, \infty]$ given by

$${}_J^\phi D(z, y) := D_J^\phi(y, z)$$

the *swapped* ϕ -selected Bregman prametric on Y . For $\partial J|_Y$ being at most single valued, we write again ${}_J^\nabla D$ and call it the *swapped Bregman prametric* on Y .

Remark 5.1.2

Please note, that there are convex functionals, whose subdifferential is single valued at a point but which is not differentiable at that same point, see e.g. [BV10, Example 4.2.6]. So in general, ∇J in our notation will not be a Gâteaux derivative.

Before going into general properties of Bregman distances and prametrics we present some examples of Bregman distances.

Example 5.1.3

- (i) Let $(H, \langle \cdot, \cdot \rangle_H)$ be a Hilbert space with induced norm $\|\cdot\|_H$ and consider $J := \frac{1}{2}\|\cdot\|_H^2$. Then J is differentiable with derivative $J'(y, z) = \langle y, z \rangle_H$ for all $y, z \in H$, and therefore $\partial J(y) = \{J'(y, \cdot)\}$ for all $y \in H$. Consequently we have

$$D_J^\nabla(z, y) = \frac{1}{2}\|z\|_H^2 - \frac{1}{2}\|y\|_H^2 - \langle y, z - y \rangle_H = \frac{1}{2}\|y - z\|_H^2.$$

So, the weighted squared norm is a Bregman distance and, since its effective domain is H , also a Bregman prametric.

- (ii) For a bounded Lebesgue measurable set $\Omega \subset \mathbb{R}^n$ consider $V := L^1(\Omega)$ and define $J : L^1(\Omega) \rightarrow]-\infty, \infty]$ by

$$J(y) = \begin{cases} \int_{\Omega} y(t) \log(y(t)) - y(t) dt & \text{if } y \geq 0 \text{ a.e., } y \log(y) \in L^1(\Omega) \\ \infty & \text{else} \end{cases}.$$

Following [Res05] we get

$$\partial J(y) = \begin{cases} \{\log(y)\} & \text{if } y \geq \varepsilon \text{ a.e. for some } \varepsilon > 0, \quad y \in L^\infty(\Omega) \\ \emptyset & \text{else} \end{cases}$$

and consequently

$$\text{dom } \partial J = \{y \in L^1(\Omega) \mid y \geq \varepsilon \text{ a.e. for some } \varepsilon > 0, \quad y \in L^\infty(\Omega)\}.$$

The induced Bregman distance on $L^1(\Omega) \times \text{dom } \partial J$ is

$$D_J^\nabla(z, y) = \begin{cases} \int_{\Omega} z(t) \log\left(\frac{z(t)}{y(t)}\right) - z(t) + y(t) dt & \text{if } z \in \text{dom } J \\ \infty & \text{else} \end{cases}.$$

and is generally called *Kullback-Leibler divergence*. We will denote the associated Bregman prametric on $Y \subseteq V$ by D_{KL} and the corresponding swapped Bregman prametric by ${}_{KL}D$.

Further examples can e.g. be found in [Pö08, Sec. 2.2].

Remark 5.1.4

Let V be a real Banach space and $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex, $Y \subseteq V$ and ϕ be a selection of ∂J . Then the ϕ -selected Bregman prametric and the swapped ϕ -selected Bregman prametric on Y are indeed well-defined prametrics on Y , i.e. they are non-negative and vanish on the diagonal of $Y \times Y$.

Having defined the main objects of this chapter, we now turn to technical properties of Bregman prametrics. In many of the proofs given in Chapter 4, we made heavy use of assumptions on finiteness of a prametric ρ and on the relation of an index set defining a bottom slice topology to $\text{dom } \rho(\cdot, y)$, $y \in Y$. So, besides being of interest in any case, knowing such effective domains are of special importance for us.

Remark 5.1.5

Let V be a real Banach space and $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex. For $y \in \text{dom } \partial J$, $z \in Z$ and $y^* \in \partial J(y)$ we have $D_J^{y^*}(z, y) < \infty$ if and only if $z \in \text{dom } J$. Therefore, for any selection ϕ of ∂J the following identities do hold for Bregman prametrics on V :

$$\begin{aligned} \text{dom } D_J^\phi &= (\text{dom } J \times \text{dom } \partial J) \cup \{(y, y) \mid y \in V \setminus \text{dom } \partial J\}, \\ \text{dom } {}_J^\phi D &= (\text{dom } \partial J \times \text{dom } J) \cup \{(y, y) \mid y \in V \setminus \text{dom } \partial J\}, \\ \text{dom } D_J^\phi(\cdot, y) &= \text{dom } {}_J^\phi D(y, \cdot) = \begin{cases} \text{dom } J & \text{if } y \in \text{dom } \partial J \\ \{y\} & \text{else} \end{cases}, \\ \text{dom } {}_J^\phi D(\cdot, y) &= \text{dom } D_J^\phi(y, \cdot) = \begin{cases} \text{dom } \partial J & \text{if } y \in \text{dom } J \\ \{y\} & \text{else} \end{cases}. \end{aligned}$$

The corresponding domains of Bregman prametrics on $Y \subseteq V$ are obtained by intersecting the above sets with $Y \times Y$ or Y respectively.

In general, Bregman prametrics do not fulfil the triangle inequality nor are they symmetric. But sometimes, the following inequality can be used as a substitute for the lacking triangle inequality.

Lemma 5.1.6 (Three term equality)

Let V be a real Banach space and $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex. For $y_1, y_2 \in V$ let $y_1^* \in \partial J(y_1)$ and $y_2^* \in \partial J(y_2)$. Then for all $z \in V$ the equation

$$D_J^{y_2^*}(z, y_2) = D_J^{y_1^*}(z, y_1) + D_J^{y_2^*}(y_1, y_2) + \langle y_1^* - y_2^*, z - y_1 \rangle$$

does hold.

In particular for every selection ϕ of ∂J on V and all $y_1, y_2 \in \text{dom } \partial J$, $z \in V$

$$D_J^\phi(z, y_2) = D_J^\phi(z, y_1) + D_J^\phi(y_1, y_2) + \langle \phi(y_1) - \phi(y_2), z - y_1 \rangle$$

does hold.

Proof: Adding

$$0 = J(y_1) - J(y_1) - \langle y_1^*, z - y_1 \rangle + \langle y_1^*, z - y_1 \rangle - \langle y_2^*, y_1 - y_1 \rangle$$

to $D_J^{y_2^*}(z, y_2)$ proves the claim. \square

Regarding terms of solutions (cf. Definition 2.1.2) induced by Bregman prametrics and regarding uniqueness of limits with respect to the sequential convergence structure $\mathcal{S}_{D_J^\phi}$, we are interested in the question, when Bregman prametrics vanish. The first lemma deals with that problem for the case of generalized Bregman distances.

Lemma 5.1.7

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex and let $y^* \in \partial J(y)$ and $z \in V$. Then the following assertions are equivalent:

- (i) $D_J^{y^*}(z, y) = 0$
- (ii) $y^* \in \partial J(z) \cap \partial J(y)$
- (iii) $J(y) - J(z) - \langle y^*, y - z \rangle = 0$

In particular $D_J^{y^*}(z, y) = 0$ implies $z \in \text{dom } \partial J$.

Proof: (i) \Rightarrow (ii) : From $D_J^{y^*}(z, y) = 0$ we get $z \in \text{dom } J$ and the identity $J(z) = J(y) + \langle y^*, z - y \rangle$. Hence for all $\tilde{z} \in V$

$$\begin{aligned} J(\tilde{z}) - J(z) - \langle y^*, \tilde{z} - z \rangle &= \\ J(\tilde{z}) - J(y) - \langle y^*, z - y \rangle - \langle y^*, \tilde{z} - z \rangle &= D_J^{y^*}(\tilde{z}, y) \geq 0 \end{aligned}$$

and therefore $y^* \in \partial J(z)$.

(ii) \Rightarrow (iii) : Let y^* be in $\partial J(z) \cap \partial J(y)$. Then

$$0 \leq D_J^{y^*}(y, z) = J(y) - J(z) - \langle y^*, y - z \rangle = -D_J^{y^*}(z, y) \leq 0.$$

(iii) \Rightarrow (i) : $D_J^{y^*}(z, y) = -(J(y) - J(z) - \langle y^*, y - z \rangle) = 0$. \square

This has immediate consequences for Bregman prametrics.

Lemma 5.1.8

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex and let ϕ be a selection of ∂J . Then the following assertions hold:

- (i) Let $y, z \in V$. Then $D_J^\phi(z, y) = 0$ if and only if either $y \in \text{dom } \partial J$ and $\phi(y) \in \partial J(z)$ or $y \notin \text{dom } \partial J$ and $z = y$.

- (ii) Let $y, z \in V$. Then ${}_J^\phi D(z, y) = 0$ if and only if either $y \in \text{dom } \partial J$ and $\phi(z) \in \partial J(y) \cap \partial J(z)$ or $y \notin \text{dom } \partial J$ and $z = y$.

Proof: Both assertions follow immediately from Definition 5.1.1 together with Lemma 5.1.7. \square

Corollary 5.1.9

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex, let ϕ be a selection of ∂J , let $\mathcal{P}_F(y)$ be an inverse problem given by an operator $F : X \rightarrow Y$ and let $y \in \text{dom } \partial J \cap Y$. Then the sets of D_J^ϕ -solutions respectively ${}_J^\phi D$ -solutions of $\mathcal{P}_F(y)$ look as follows:

$$\mathcal{L}_F^{ex, D_J^\phi}(y) = \{x \in X \mid D_J^\phi(Fx, y) = 0\} = \{x \in X \mid \phi(y) \in \partial J(Fx)\}$$

and

$$\mathcal{L}_F^{ex, {}_J^\phi D}(y) = \{x \in X \mid Fx \in \text{dom } \partial J \text{ and } \phi(Fx) \in \partial J(y)\}.$$

Combining this with Definition 2.1.2(i), we get the following representation of solution sets with respect to Bregman prametrics.

Lemma 5.1.10

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex, let ϕ be a selection of ∂J and $Y \subseteq V$. Then the following statements are equivalent:

- (i) The ϕ -selected Bregman prametric D_J^ϕ on Y is separating.
- (ii) The swapped ϕ -selected Bregman prametric ${}_J^\phi D$ on Y is separating.
- (iii) For all $y \in Y \cap \text{dom } \partial J$ we have

$$\phi(y) \in \bigcap_{z \in Y \setminus \{y\}} (V^* \setminus \partial J(z)). \quad (5.1)$$

Proof: Equivalence of (i) and (ii) is due to $D_J^\phi(z, y) = {}_J^\phi D(y, z)$. Therefore it is sufficient to show (i) \Leftrightarrow (iii). Let ϕ be a selection of ∂J . Since D_J^ϕ is separating on $V \setminus \text{dom } \partial J$ in any case, we only need to examine $Y \cap \text{dom } \partial J$. Let D_J^ϕ be separating and $y, z \in Y$. Due to Lemma 5.1.8(i) we get $\phi(y) \in \partial J(z)$ if and only if $z = y$. Therefore, $\phi(y) \notin \partial J(z)$ for all $z \in Y \setminus \{y\}$. Conversely, if $\phi(y) \notin \partial J(z)$ for all $z \in Y \setminus \{y\}$, then Lemma 5.1.8(i) implies $D_J^\phi(z, y) \neq 0$ for all $z \neq y$. \square

As it turns out, strictly convex functionals J have a particular nice behaviour in that context.

Lemma 5.1.11

Let V be a real Banach space and let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and strictly convex. Then D_J^ϕ and ${}_J^\phi D$ are separating for all selections ϕ of ∂J .

Proof: Since J is strictly convex, the subdifferential ∂J is strictly monotone, i.e. $\langle y^* - z^*, y - z \rangle > 0$ for all $y \neq z \in Y$ and $y^* \in \partial J(y)$, $z^* \in \partial J(z)$. Therefore $\partial J(y) \cap \partial J(z) = \emptyset$ for all $y \neq z \in Y$ and consequently every selection ϕ of ∂J satisfies the inclusion (5.1). \square

It would be nice, if we could relate being separated of (swapped) ϕ -selected Bregman prametrics to mapping properties of ϕ as e.g. injectivity. In the case, that all the subdifferentials of elements of Y are singletons or empty this is indeed possible.

Corollary 5.1.12 (see also [LW13])

Let V be a real Banach space, $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ and $Y \subseteq V$ such that $(\partial J)|_Y$ is at most single valued. Then the following assertions hold:

- (i) Let $y \in Y \cap \text{dom } \partial J$ and $z \in Y$. Then $D_J^\nabla(z, y) = 0$ if and only if $z \in Y \cap \text{dom } \partial J$ and $\nabla J(z) = \nabla J(y)$.
- (ii) Let $y \in Y \cap \text{dom } \partial J$ and $z \in Y$. Then ${}_J^\nabla D(y, z) = 0$ if and only if $z \in Y \cap \text{dom } \partial J$ and $\nabla J(z) = \nabla J(y)$.
- (iii) The following statements are equivalent:
 - (a) The Bregman prametric D_J^∇ on Y is separating.
 - (b) The swapped Bregman prametric ${}_J^\nabla D$ on Y is separating.
 - (c) $\nabla J : Y \cap \text{dom } \partial J \rightarrow V^*$ is injective.

Proof: Items (i) and (ii) follow immediately from Lemma 5.1.8. Item (iii) is a consequence of (i) and Lemma 5.1.10. \square

Again, strictly convex functionals are especially well-behaved.

Corollary 5.1.13 (see also [LW13])

Let V be a real Banach space, $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ such that ∂J is at most single valued. If J is strictly convex, then D_J^∇ on V is separating. If J is Gâteaux differentiable then the converse also holds.

Proof: The first assertion is simply Lemma 5.1.11. Now let J be Gâteaux differentiable and let D_J^∇ on V be separating. Then, due to Lemma 5.1.12, ∇J is injective. Suppose now, that J is not strictly convex. Since J is differentiable, ∇J is not strictly monotone (due to [Sch07, Prop. 4.3.5]).

Consequently there exist $y \neq z \in V$ such that $J(y) - J(z) = \langle \nabla J(z), y - z \rangle$. This implies for arbitrary $z' \in V$, that

$$J(z') - J(y) - \langle \nabla J(z), z' - y \rangle = J(z') - J(z) - \langle \nabla J, z' - z \rangle = D_J^\nabla(z', z) \geq 0.$$

Hence we have $\nabla J(z) \in \partial J(y) = \{\nabla J(y)\}$ in contradiction to ∇J being injective. \square

5.2 Bregman prametrics, topological issues and Tikhonov regularization

In this section we take a closer look on the topologies and sequential convergence structures constructed from prametrics for the special case of Bregman prametrics. First we have two short remarks on prametric topologies and the sequential convergence structure induced by a Bregman prametric.

Remark 5.2.1

Let V be a Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be convex and proper, let $Y \subseteq V$ and let ϕ be a selection of ∂J . Denote by $\tau_{D_J^\phi}$ and $\tau_{\phi D}$ the prametric topologies with respect to D_J^ϕ and ϕD . Then the following assertions hold:

- (i) Every subset $U \subseteq (V \setminus \text{dom } J) \cap Y$ is $\tau_{D_J^\phi}$ -open and $\text{dom } J \cap Y$ is $\tau_{D_J^\phi}$ -closed.
- (ii) Every subset $U \subseteq (V \setminus \text{dom } \partial J) \cap Y$ is $\tau_{\phi D}$ -open and $\text{dom } \partial J \cap Y$ is $\tau_{\phi D}$ -closed.

Proof:

- (i) Let $y_0 \in U \subseteq (V \setminus \text{dom } J) \cap Y$. Since $D_J^\phi(y_0, y) = \infty$ for all $y \neq y_0$, we have $\mathcal{B}_\varepsilon^{D_J^\phi}(y_0) = \{y_0\} \subseteq U$ for all $0 < \varepsilon < \infty$. Hence, U is $\tau_{D_J^\phi}$ -open. In particular, $V \setminus \text{dom } J$ is itself open, and therefore $\text{dom } J$ is open.
- (ii) Goes analogously to (i).

\square

Remark 5.2.2

Let V be a Banach space, $J : V \rightarrow [0, \infty]$ be proper and convex, ϕ a selection of $\text{dom } \partial J$ and $Y \subseteq \text{dom } \partial J$. If the sequential convergence structure $\mathcal{S}_{D_J^\phi}$ fulfils [CONT], then it has unique limits if and only if D_J^ϕ is separable on Y .

Proof: Since $y \in \text{dom } \partial J$ for all $y \in Y$, we have $D_J^\phi(z, y) < \infty$ for all $z, y \in Y$. In particular, this does hold for all $z, y \in Y$ such that $\mathcal{S}_{D_J^\phi}(z) \cap \mathcal{S}_{D_J^\phi} \neq \emptyset$. So, if we have a sequence (y_n) converging to both, z and y , [CONT] implies $D_J^\phi(z, y_n) \rightarrow D_J^\phi(z, y)$ and Lemma 1.3.19 proves the claim. \square

Now we take a turn to sequential convergence with respect to bottom slice topologies induced by prametrics. The results concerning this topic are in essence generalizations of assertions stated in [LW13, Sec. 3.3] to the case of (swapped) ϕ -selected Bregman prametrics.

Since the amount of sub- and superscripts in the notion introduced before would now definitively get unbearable, we will use a somewhat simplified notation.

Notation

Let V be a Banach space and $J : V \rightarrow [0, \infty]$ be proper and convex, ϕ a selection of $\text{dom } \partial J$ and $\tilde{Y}, Z \subseteq Y \subseteq V$. Then we denote by

- $\tau_{J, \phi}^Z := \tau_{IN}^{Z, D_J^\phi}$ the bottom slice topology with respect to Z and the ϕ -selected Bregman prametric D_J^ϕ on \tilde{Y}
- $\tau_{\phi, J}^{Z, \phi} := \tau_{IN}^{Z, \phi D}$ the bottom slice topology with respect to Z and the swapped ϕ -selected Bregman prametric ${}^\phi D$ on \tilde{Y}

Now, we have the following characterization of ϕ -selected Bregman prametrics.

Lemma 5.2.3

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex, let ϕ be a selection of ∂J and let $\tilde{Y}, Z \subseteq Y \subseteq V$. By $\tau_{J, \phi}^Z$ we denote the bottom slice topology with respect to D_J^ϕ on \tilde{Y} .

Consider $y \in \tilde{Y} \cap Z \cap \text{dom } \partial J$ and a sequence (y_n) being eventually in $\tilde{Y} \cap \text{dom } \partial J$. Then, the following statements are equivalent:

- (i) $y_n \xrightarrow{\tau_{J, \phi}^Z} y$
- (ii) $D_J^\phi(z, y_n) \rightarrow D_J^\phi(z, y)$ for all $z \in Z$
- (iii) $D_J^\phi(y, y_n) \rightarrow 0$ and $\langle \phi(y_n) - \phi(y), y - z \rangle \rightarrow 0$ for all $z \in Z \cap \text{dom } J$.
- (iv) $D_J^\phi(y, y_n) \rightarrow 0$ and $\phi(y_n) \xrightarrow{*} \phi(y)$ in $\text{span}(y - (Z \cap \text{dom } J))^*$.

Proof: Equivalence of (i) and (ii) is a specialization of Lemma 4.2.7.

Next we show (ii) \Leftrightarrow (iii). Let $z \in \text{dom } J$. By adding $0 = \langle \phi(y_n), y - y \rangle$ we get

$$D_J^\phi(z, y_n) - D_J^\phi(z, y) = D_J^\phi(y, y_n) + \langle \phi(y_n) - \phi(y), y - z \rangle \quad (5.2)$$

for all $n \in \mathbb{N}$ large enough.

Now let $D_J^\phi(z, y_n) \rightarrow D_J^\phi(z, y)$ hold for all $z \in Z$. Due to $y \in Z$ this implies $D_J^\phi(y, y_n) \rightarrow D_J^\phi(y, y) = 0$ and the first assertion of (iii) is shown. Passing on to the limit in (5.2) shows the second assertion $\langle \phi(y_n) - \phi(y), y - z \rangle \rightarrow 0$.

Conversely let (iii) hold. Consider first $z \notin \text{dom } J$. Due to Remark 5.1.5 we have $D_J^\phi(z, y) = \infty$ and $D_J^\phi(z, y_n) = \infty$ for all $n \in \mathbb{N}$ large enough and therefore $D_J^\phi(z, y_n) \rightarrow D_J^\phi(z, y)$. Therefore, considering for $z \in \text{dom } J$ the limit of equation (5.2) proves (ii).

For proving the last equivalence, let $\tilde{Z} := y - (Z \cap \text{dom } J)$. First, let (iii) hold. Since $\text{span}(\tilde{Z})$ is a subspace of V we have $V^* \subseteq \text{span}(\tilde{Z})^*$, therefore $\phi(y), \phi(y_n) \in \text{span}(\tilde{Z})^*$.

Now consider arbitrary $\tilde{z} \in \text{span}(\tilde{Z})$. Then there exist $\beta_1, \dots, \beta_k \in \mathbb{R}$ and $z_1, \dots, z_n \in Z \cap \text{dom } J$ such that $\tilde{z} = \sum_{j=1}^k \beta_j(y - z_j)$. Since $\phi(y) - \phi(y_n)$ is a linear mapping, this implies

$$\langle \phi(y_n) - \phi(y), \tilde{z} \rangle = \sum_{j=1}^n \beta_j \langle \phi(y_n) - \phi(y), y - z_j \rangle \rightarrow 0$$

which proves (iv). The converse does hold since $y - z \in \text{span}(\tilde{Z})$ for all $z \in Z \cap \text{dom } J$. \square

Corollary 5.2.4

Let V be a Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be convex and proper, let $Y \subseteq V$, ϕ a selection of ∂J and $\tilde{Y} \subseteq Y \cap \text{dom } \partial J$ and $Z \subseteq Y \cap \text{dom } J$ such that $0 \in Z$. By $\tau_{J, \phi}^Z$ we denote the bottom slice topology with respect to D_J^ϕ and Z on \tilde{Y} .

- (i) If $y \in \tilde{Y} \cap Z$ and (y_n) is a sequence in \tilde{Y} , then $y_n \xrightarrow{\tau_{J, \phi}^Z} y$ if and only if $D_J^\phi(y, y_n) \rightarrow 0$ and $\phi(y_n) \xrightarrow{*} \phi(y)$ in $\text{span}(Z)^*$
- (ii) If $\tilde{Y} \subseteq Z$ does hold, then the mapping $\tilde{\phi} : \tilde{Y} \rightarrow \text{span}(Z)^*$ given by $\tilde{\phi}(y) := \phi(y)$ is sequentially $\tau_{J, \phi}^Z$ -weak*-continuous.

Proof:

- (i) Due to $0 \in Z$ we have $z = (y - 0) - (y - z) \in \text{span}(y - Z)$ for all $z \in Z$. Conversely, we have $y - z \in \text{span}(Z)$ because of $y \in \tilde{Y} \cap Z$. So, we have $\text{span}(Z) = \text{span}(y - Z) = \text{span}(y - (Z \cap \text{dom } J))$ and the assertions follows from Lemma 5.2.3.

(ii) Is a direct consequence of (i).

□

Corollary 5.2.5 ([LW13])

Let V be a Banach space and $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be convex and finite, let ϕ be a selection of ∂J and set $\tilde{Y} := \text{dom } \partial J$.

If the prametric topology $\tau_{D_J^\phi}$ on \tilde{Y} fulfils [CONT], then ϕ is $\tau_{D_J^\phi}$ -weak * continuous.

Proof: Since $V = \text{dom } J = \text{dom } D_J^\phi(\cdot, \tilde{y})$ for all $\tilde{y} \in \tilde{Y}$, Theorem 4.2.13 implies

$$\mathcal{S}(\tau_{J,\phi}^Z) = \mathcal{S}(\tau_{D_J^\phi}) = \mathcal{S}_{(D_J^\phi)|_{\tilde{Y} \times \tilde{Y}}}.$$

Since $\tau_{D_J^\phi}$ is sequential and due to Corollary 5.2.4 this proves the claim. □

Sadly, such a nice interpretation of convergence with respect to the bottom slice topology comes not as readily for the swapped ϕ -selected Bregman prametric. But at least we have the following characterization of convergence in $\tilde{Y} \cap \text{dom } \partial J$.

Lemma 5.2.6

Let V be a real Banach space, let $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ be proper and convex, let ϕ be a selection of ∂J and let $\tilde{Y}, Z \subseteq Y \subseteq V$. By $\phi_{\cdot,J}^Z \tau$ we denote the bottom slice topology with respect to ${}^\phi J D$ on \tilde{Y} .

Consider $y \in \tilde{Y} \cap Z \cap \text{dom } \partial J$ and a sequence (y_n) being eventually in $\tilde{Y} \cap \text{dom } \partial J$. Then, the following statements are equivalent:

- (i) $y_n \xrightarrow{\phi_{\cdot,J}^Z \tau} y$.
- (ii) $D_J^\phi(y_n, z) \rightarrow D_J^\phi(y, z)$ for all $z \in Z$.
- (iii) ${}^\phi J D(y, y_n) = D_J^\phi(y_n, y) \rightarrow 0$ and $\langle \phi(y) - \phi(z), y - y_n \rangle \rightarrow 0$ for all $z \in Z \cap \text{dom } \partial J$.

Proof: Due to Lemma 4.2.7 and ${}^\phi J D(z, y) = D_J^\phi(z, y)$ we have (i) \Leftrightarrow (ii).

For the second equivalence let $z \in Z \cap \text{dom } \partial J$. Due to $y \in \text{dom } \partial J$ and $y_n \in \text{dom } J$, adding $0 = \langle \phi(y) - \phi(y), y_n - y \rangle$ leads to

$$D_J^\phi(y_n, z) - D_J^\phi(y, z) = D_J^\phi(y_n, y) - \langle \phi(y) - \phi(z), y_n - y \rangle.$$

for $n \in \mathbb{N}$ large enough. Using $y \in Z$ this directly implies (iii) \Rightarrow (ii) and the converse for $z \in \text{dom } \partial J$. Now let $z \notin \text{dom } \partial J$. Since $y_n \in \text{dom } \partial J$ for all $n \in \mathbb{N}$ large enough, we have $D_J^\phi(y_n, z) = D_J^\phi(y, z) = \infty$ for all n large enough. □

Remark 5.2.7

For all implications except (iii) \Rightarrow (ii) in the case of $z \notin \text{dom } \partial J$, it would be sufficient to assume (y_n) eventually in $\text{dom } J$. The problem in that last case is, that $y_n = z \in \text{dom } J \setminus \text{dom } \partial J$ for infinitely many $n \in \mathbb{N}$ (and consequently $D_J^\phi(y_n, z) = 0$ but $D_J^\phi(y, z) = \infty$ for infinitely many n) could happen. If $Z \subseteq \text{dom } \partial J$ does hold, the assumption (y_n) eventually in $\text{dom } \partial J$ is superfluous. The same is valid, if $y_n \notin Z$ is known for all n large enough.

As a closing to this considerations, we present a concrete example for the consequences of the findings in Lemma 5.2.3 taken from [LW13].

Example 5.2.8 ([LW13])

In this example we consider the Bregman parametric D_{KL} constructed from the Kullback-Leibler divergence defined in Example 5.1.3 as in Definition 5.1.1. Set $Z := \text{dom } J = \{y \in L^1(\Omega) \mid y \geq 0 \text{ a.e., } y \log(y) \in L^1(\Omega)\}$ and $\tilde{Y} := \text{dom } \partial J = \{y \in L^1(\Omega) \cap L^\infty(\Omega) \mid y \geq \varepsilon \text{ a.e. for some } \varepsilon > 0\}$.

According to [LW13, Lemma 3.23] we have $\text{span}(Z)^* = L^\infty(\Omega)$. Using this, Corollary 5.2.4 and the fact, that $D_{KL}(y, y_n) \rightarrow 0$ implies convergence with respect to the L^1 -norm, [LW13, Thm. 3.24] can be proven, which states that on \tilde{Y} we have

$$y_n \xrightarrow{\tau_{TN}^{Z, D_{KL}}} y \Leftrightarrow \begin{cases} y_n \rightarrow y & \text{in } L^1(\Omega) \text{ and} \\ \log(y_n) \xrightarrow{*} \log(y) & \text{in } L^\infty(\Omega) \end{cases}.$$

In Chapter 4 we already mentioned, that bottom slice topologies are themselves candidates for a topology inducing the sequential convergence structure of a regularizing Tikhonov setup. So, we end this thesis by presenting a setting where a bottom slice topology actually appears in such a role.

Notation

Let X be a real Banach space. Then we denote by τ_w the weak topology on X and we write $x_n \rightharpoonup x$ if (x_n) converges weakly to x .

We will use a ϕ -selected Bregman parametric as discrepancy functional. In Section 3.3 we have seen, that it is necessary for a variational setup to be a regularizing Tikhonov setup, that the domain of the regularization functional intersected with the preimage of the domain of all $D_J^\phi(\cdot, y)$, $y \in Y$ is non-empty. Looking at Remark 5.1.5, allowing for $\text{dom } \partial J \subsetneq Y$ would imply, that we had to assume $Y \setminus \text{dom } \partial J \subset \text{rg } F$ for the operator defining the inverse problem of such a variational setup. Therefore we decided only to cover the case $Y = \text{dom } \partial J$.

General assumption 5.2.9

In the Rest of this chapter, we assume that

-
- (i) $(X, \|\cdot\|_X)$ is a reflexive Banach space
 - (ii) $R : X \rightarrow [0, \infty]$ is weakly sequentially lower semi-continuous and coercive, i.e. $\|x_n\|_X \rightarrow \infty$ implies $R(x_n) \rightarrow \infty$
 - (iii) V is a real Banach space, $J : V \rightarrow \mathbb{R} \cup \{\infty\}$ is proper and convex and ϕ is a selection of ∂J
 - (iv) $\tau_{J,\phi}^Z$ is the bottom slice topology with respect to D_J^ϕ and $Z := \text{dom } J$ on $Y := \text{dom } \partial J$
 - (v) F is sequentially weak- $\tau_{J,\phi}^Z$ continuous
 - (vi) D_J^ϕ is sequentially $\tau_w \times \tau_{J,\phi}^Z$ -lower semi-continuous
 - (vii) $\text{dom } R \cap F^{-1}(Y) \neq \emptyset$

Theorem 5.2.10

$\mathcal{M} := ((X, \tau_w), (Y, \mathcal{S}(\tau_{J,\phi}^Z), \tau_{J,\phi}^Z), \mathcal{P}_F, (D_J^\phi)_{|Y \times Y}, R)$ is a regularizing Tikhonov setup.

Proof: Since X is reflexive and since R is weakly sequentially lower semi-continuous, the sublevelsets of R are weakly sequentially compact. Furthermore we have $Y \subseteq Z$ and therefore $y_n \xrightarrow{\tau_{J,\phi}^Z} y$ implies $D_J^\phi(z, y_n) \rightarrow D_J^\phi(z, y)$ for all $z \in \text{dom } D_J^\phi(\cdot, y) \cap Y = Y$. Consequently, the proof of Theorem 4.1.2 applies literally to this case. \square

Looking at this theorem, it would be interesting to go on with the technical investigations from the previous section, by studying what item (v) and (vi) of our assumptions actually mean and in particular, if we do not talk about the empty set. But this is also left for future work.

Conclusion

In the first two chapters, we gave axiomatic definitions for a lot of terms being central in dealing with inverse problems. This included sequential convergence, stability, regularization methods and in particular, terms of solutions. Although serving their purpose in this thesis, all these definitions are certainly not the last word on that subject, but it seems rather questionable if there is 'the right definition' for these objects at all. As e.g. illustrated by Remark 2.1.4 and by Remark 2.1.6, a final judgement on what is really sensible in a real world context depends for one thing on the modelling of the underlying material problem and for another on the aim being pursued by solving such a problem. At the end of the day, these are issues which cannot be decided on a purely inner-mathematical basis.

Nevertheless, mathematics is far from being useless in the context of inverse problems. Besides interdisciplinary work on the modelling of real world problems, two of the main things mathematicians can add to inverse problems are the following.

First, they can provide concrete methods and theory for the solution of particular (classes of) inverse problems with a 'clear' modelling, adapt and optimize them to the needs of special cases and study their behaviour, as e.g. error bounds and convergence speed and contribute thereby directly to the solution of practical inverse problems. These are things, this thesis does certainly not accomplish to a noteworthy extent.

Second, it is likewise useful and important to work on a more theoretical level having no specific application in mind, for example searching the toolkit of mathematics for instruments being possibly helpful in the field of inverse problems or analyzing already existent, good working theory and techniques, both providing impulses for possible ways in future work. Thereby the latter point includes checking, what is actually covered by such tools in their current form, where they can be extended to which more general settings, where entirely different techniques would be required and what is going to work mathematically in principle independent of its current usability. To this, this work is a contribution.

Starting from a list of questions in the introduction, it presented sequential convergence spaces as an alternative frame for inverse problems and studied some properties of prametrics concerning their adequacy as similarity measure in Tikhonov regularization and their fitness for already existing theoretical frameworks. It proposed definitions of well-posedness, stability and regularization methods, which are able to incorporate this 'new' objects and which additionally reflect the dependence of these terms on what is actually regarded as a solution of an inverse problem. Moreover, it explored two ways of constructing topologies from prametrics, both of them possessing properties desirable for Tikhonov regularization and used them to determine when a common set of sufficient conditions can be situated in a purely topological framework (and besides demonstrated, that such a purely topological framework is not necessary for the used techniques to work). Finally, it used some of the developed concepts to study useful properties of a special class of prametrics deduced from Bregman distances.

Obviously, this thesis is far from exhausting this topics. As pointed out in Chapter 4, the restriction to prametrics in the definition of bottom slice topologies is somewhat limiting and it would be an interesting question to study models with separated spaces for exact and measured data as by Fleming using analogue constructions. Really using the specific construction of Tikhonov regularization methods could lead to necessary conditions on regularizing Tikhonov setups being more meaningful than the ones given in Chapter 3. Moreover, continuing the study of prametric topologies, sequential convergence spaces and bottom slice topologies induced by Bregman prametrics started in Chapter 5 by explicitly calculating these objects and studying their properties as e.g. various sorts of continuity may give deeper insight in this special class of discrepancy functionals and may make them easier accessible to regularization theory. Finally, a deeper study of the utility of sequential convergence spaces and convergence spaces in the context of inverse problems might be desirable, since after [BB02] some of them provide parallels to important results and concepts from functional analysis.

By this incomplete list of possible future work regarding theoretical details of the thesis as well as the accessibility of the introduced objects and instruments to the theory of inverse problems, we end this thesis.

Notation

$\phi, J^{\mathcal{Z}} \tau$	bottom slice topology w.r.t a swapped ϕ -selected Bregman prametric, page 67
2^X	power set of the set X , page 5
A^*	adjoint operator, page 26
$\mathcal{B}_\varepsilon^\rho(y_0)$	ε -ball w.r.t. ρ centered at y_0 , page 13
D_J^ϕ, D_J^∇	, page 60
${}^\phi J D, {}^\nabla J D$	swapped Bregman prametric, page 60
$\tau_{IN}, \tau_{IN}^{\mathcal{Z}, \rho}$, page 51
$\tau_{J, \phi}^{\mathcal{Z}}$	bottom slice topology w.r.t a ϕ -selected Bregman prametric, page 67
$\mathcal{C}_s(X, Y)$	continuous mappings between sequential convergence spaces, page 6
$\mathcal{D}(F)$	domain of definition of a partial function F , page 23
$\text{dom}(\mathcal{F})$	domain of a set-valued mapping, page 24
eq	, page 12
$\text{gr}(\mathcal{F})$	graph of a set-valued mapping, page 24
$\langle \cdot, \cdot \rangle$	dual pairing, page 59
\mathcal{P}_F	inverse problem given by F , page 23
$\mathcal{P}_F(y)$, page 23
\mathcal{L}^\perp	solution term w.r.t normal equation, page 26
\mathcal{L}	term of solution, page 24

\mathcal{L}^{\min, x^*}	x^* -minimum-norm-solution non-linear case, page 27
$\mathcal{L}^{\dagger, x^*}$	solution term w.r.t x^* -minimum-norm-solution, page 26
$\mathcal{L}^{\min R, \rho}$	R -minimum ρ -solution, page 27
A^\dagger	Moore-Penrose inverse, page 26
\mathcal{L}^\dagger	solution term w.r.t Moore-Penrose inverse, page 26
$\text{Map}_P(X, Y)$	partial mappings from X to Y , page 24
π_M	projection onto M , page 26
\mathbb{R}^+	positive real numbers, page 34
$\text{rg}(F)$	range of a mapping F , page 26
$\tau _M$	subspace topology on M , page 2
$\mathcal{S}(\tau)$	sequential convergence structure induced by a topology τ , page 6
\mathcal{S}_ρ	sequential convergence structure induced by a prametric ρ , page 17
$\mathcal{L}_F^{ex}(y)$	set of exact solutions of $\mathcal{P}_F(y)$, page 24
$\mathcal{L}_F^{ex, \rho}(y)$	set of ρ -solutions, page 24
\mathcal{L}_F	, page 25
∂J	subdifferential of J , page 59
$\tau(\mathcal{S})$	the topology induced by a sequential convergence structure \mathcal{S} , page 9
τ_ρ	prametric topology w.r.t. ρ , page 13
$T_{\alpha, y}(x)$	Tikhonov functional, page 39
\mathcal{T}	Tikhonov operator, page 39
$\text{dom}(f)$	effective domain of an extended real-valued mapping, page 24
$L(X, Y)$	set of linear bounded operators between two normed spaces X and Y , page 26
M^\perp	orthogonal complement of a set M , page 26
V^*	topological dual of a normed vector space V , page 59

$x_n \rightharpoonup x$ weak convergence, page 70

$y_n^* \xrightarrow{*} y^*$ weak* convergence, page 59

[CONT] , page 48

[CONV] , page 48

Index

- ϕ -selected Bregman prametric, 60
- ρ -solution, 24
- \mathcal{S} -limit point, 9
- τ - \mathcal{S} -continuous, 7
- x^* -minimum-norm-solution, 27
- [CONT], 48, 50
- [CONV], 48, 50
- R -minimum ρ -solution, 27
- bottom slice topology, 51
- Bregman distance, 60
- Bregman prametric, 60
 - ϕ -selected, 60
 - swapped, 60
 - swapped ϕ -selected, 60
- category, well-mannered, 30
- closed
 - sequentially, 1
- compact
 - sequentially, 1
- continuity
 - $\tau - \mathcal{S}, \mathcal{S} - \tau$, 7
 - sequential convergence spaces, 6
 - well-mannered categories, 30
- convergence
 - in sequential convergence spaces, 5
 - stronger, 6
 - weaker, 6
- discrepancy functional, 39
- eventually in U , 1
- exact solution, 24
- exact term of solution, 24
- generalized Bregman distance, 60
- globally ill-posed, 31
- globally well-posed, 31
- graph subconvergent, 7
- ill-posed, 31
 - globally, 31
 - locally, 32
- initial topology, 3
- inverse problem, v, 23
- Kullback-Leibler distance, 61
- open
 - sequentially, 1
- order topology, 5
- prametric, 11
 - separating, 11
- prametric topology, 13
- regularization functional, 39
- regularization method, 36
- regularization operator, 36
- regularizing Tikhonov setup, 42
- sc-open, 9
- sequential convergence space, 5
 - topological, 9
- sequential convergence structure, 5

- induced by a prametric, 17
- induced by a topology, 6
- stronger, 6
- topological, 9
- weaker, 6
- sequential topology, 1
- sequentially closed, 1
- sequentially compact, 1
- sequentially open, 1
- set of ρ -solutions, 24
- set of exact solutions, 24
- solution, 24
- stability of an inverse problem, 32
- stability of set-valued mappings, 32
- subconvergent, 7
- subdifferential, 59
- subsequential convergence, 8
- subsequential limit, 8
- swapped Bregman prametric, 60
- term of solution, 24
- Tikhonov functional, 39
- Tikhonov operator, 39
- topology
 - induced by a sequential convergence structure, 9
 - initial, 3
 - prametric, 13
 - sequential, 1
- variational setup, 41
- well-mannered category, 30
- well-posed, 31
 - globally, 31
 - locally, 32

Bibliography

- [AF90] Alexander V. Arkhangel'skiĭ and Vitaly V. Fedorchuk. Basic concepts and constructions of general topology. In L. S. Pontryagin, editor, *General topology. I*, Encyclopaedia of Mathematical Sciences. Springer, 1990.
- [AV94] Robert Acar and Curtis R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10(6):1217–1229, 1994. ausgedruckt.
- [Bak84] Anatoly B. Bakushinskiĭ. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *Akademiya Nauk SSSR. Zhurnal Vychislitel' noi Matematiki i Matematicheskoi Fiziki*, 24(8):1258–1259, 1984.
- [Bau68] Heinz Bauer. *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*. Walter de Gruyter & Co., Berlin, 1968.
- [BB02] Ronald Beattie and Heinz-Peter Butzmann. *Convergence structures and applications to functional analysis*. Kluwer Academic Publishers, Dordrecht, 2002.
- [BB11] Martin Benning and Martin Burger. Error Estimates for General Fidelities . *Electronic Transactions on Numerical Analysis*, 38:44–68, March 2011.
- [BIG03] Adi Ben-Israel and Thomas N. E. Greville. *Generalized inverses*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 15. Springer-Verlag, New York, second edition, 2003. Theory and applications.
- [BL08] Kristian Bredies and Dirk A. Lorenz. Linear convergence of iterative soft-thresholding. *J. Fourier Anal. Appl.*, 14(5-6):813–837, 2008. ausgedruckt.

- [BO04] Martin Burger and Stanley Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411–1421, 2004. ausgedruckt.
- [Bou66] Nicolas Bourbaki. *Elements of mathematics. General topology. Part 1*. Hermann, Paris, 1966.
- [Buz08] Thorsten M. Buzug. *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer-Verlag, Berlin/Heidelberg, 2008.
- [BV10] Jonathan M. Borwein and Jon D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Cambridge University Press, New York, 2010.
- [CK94] Guy Chavent and Karl Kunisch. Convergence of Tikhonov regularization for constrained ill-posed inverse problems. *Inverse Problems*, 10(1):63–76, 1994.
- [DDD04] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11):1413–1457, 2004. ausgedruckt.
- [EHN96] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [Eng81] Heinz W. Engl. Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind. *Numer. Funct. Anal. Optim.*, 3(2):201–222, 1981.
- [Fle10] Jens Flemming. Theory and examples of variational regularization with non-metric fitting functionals. *J. Inverse Ill-Posed Probl.*, 18(6):677–699, 2010.
- [Fle11] Jens Flemming. *Generalized Tikhonov regularization*. PhD thesis, Technische Universität Chemnitz, Chemnitz, may 2011.
- [Fra65] Stanley P. Franklin. Spaces in which sequences suffice. *Fund. Math.*, 57:107–115, 1965.

-
- [Fra67] Stanley P. Franklin. Spaces in which sequences suffice. II. *Fund. Math.*, 61:51–56, 1967.
- [Fre06] Maurice Frechet. Sur quelques point du calcul fonctionnel. *Rend. Circ. Mat. Palermo*, 22(1):1–72, 1906.
- [GHS11a] Markus Grasmair, Markus Haltmeier, and Otmar Scherzer. Necessary and sufficient conditions for linear convergence of ℓ^1 -regularization. *Comm. Pure Appl. Math.*, 64(2):161–182, 2011.
- [GHS11b] Markus Grasmair, Markus Haltmeier, and Otmar Scherzer. The residual method for regularizing ill-posed problems. *Applied Mathematics and Computation*, 218(6):2693–2710, 2011.
- [Gor04] Anthony Goreham. Sequential Convergence in Topological Spaces. arXiv:math/0412558v1 [math.GN], December 2004. Survey, Preprint.
- [Had02] Jaques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, page 49–52, 1902.
- [HKPS07] Bernd Hofmann, Barbara Kaltenbacher, Christiane Pöschl, and Otmar Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23(3):987–1010, 2007. ausgedruckt.
- [Hof00] Bernd Hofmann. Ill-posedness and local il-posedness concepts in hilbert spaces. *Optimization*, 48(2):219–238, 2000.
- [HW13] Thorsten Hohage and Frank Werner. Iteratively regularized Newton-type methods for general data misfit functionals and applications to Poisson data. *Numer. Math.*, 123(4):745–779, April 2013.
- [Iva69] Valentin Konstantinovich Ivanov. Incorrect problems in topological spaces. *Akademija Nauk SSSR. Sibirskoe Otdelenie. Sibirskii Matematičeskii Žurnal*, 10:1065–1074, 1969.
- [Kis60] Jan Kiszyński. Convergence du type \mathcal{L} . *Colloq. Math.*, 7:205–211, 1959/1960.
- [KS05] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.

- [LW13] Dirk Lorenz and Nadja Worliczek. Necessary conditions for variational regularization schemes. *Inverse Problems*, 29(7):075016, 2013.
- [Mun00] James R. Munkres. *Topology*. Prentice Hall, Incorporated, second edition, 2000.
- [Nas71] M. Zuhair Nashed. Generalized inverses, normal solvability, and iteration for singular operator equations. In *Nonlinear Functional Anal. and Appl. (Proc. Advanced Sem., Math. Res. Center, Univ. of Wisconsin, Madison, Wis., 1970)*, page 311–359. Academic Press, New York, 1971.
- [Nat86] Frank Natterer. *The mathematics of computerized tomography*. B. G. Teubner, Stuttgart, 1986.
- [Ord66] Edward T. Ordman. Classroom Notes: Convergence Almost Everywhere is Not Topological. *Amer. Math. Monthly*, 73(2):182–183, 1966.
- [Pö08] Christiane Pöschl. *Tikhonov Regularization with General Residual Term*. PhD thesis, Leopold Franzens Universität Innsbruck, Innsbruck, October 2008.
- [RA07] Elena Resmerita and Robert S. Anderssen. Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems. *Math. Methods Appl. Sci.*, 30(13):1527–1544, 2007. ausgedruckt.
- [Res05] Elena Resmerita. Regularization of ill-posed problems in Banach spaces: convergence rates. *Inverse Problems*, 21(4):1303–1314, 2005. ausgedruckt.
- [Rie03] Andreas Rieder. *Keine Probleme mit inversen Problemen*. Friedr. Vieweg & Sohn, Braunschweig, 2003. Eine Einführung in ihre stabile Lösung. [An introduction to their stable solution].
- [RO94] Leonid I. Rudin and Stanley Osher. Total Variation Based Image Restoration with Free Local Constraints. In *ICIP (1)*, page 31–35. IEEE, 1994.
- [ROF92] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, November 1992.

-
- [Sch07] Winfried Schirotzek. *Nonsmooth analysis*. Universitext. Springer, Berlin, 2007.
- [Sei81] Thomas I. Seidman. Convergent approximation methods for ill-posed problems. I. General theory. *Control Cybernet.*, 10(1-2):31–49 (1982), 1981.
- [SGG⁺09] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.
- [SKHK12] Thomas Schuster, Barbara Kaltenbacher, Bernd Hofmann, and Kamil S. Kazimierski. *Regularization methods in Banach spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2012.
- [TA77] Andrey N. Tikhonov and Vasily Y. Arsenin. *Solutions of Ill-posed Problems*. V.H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977.
- [Tre10] Dennis Trede. *Inverse problems with sparsity constraints: Convergence rates and exact recovery*. PhD thesis, Universität Bremen, 2010.
- [Tro06] Joel A. Tropp. Just relax: convex programming methods for identifying sparse signals in noise. *IEEE Trans. Inform. Theory*, 52(3):1030–1051, 2006. ausgedruckt.
- [Ury26] Pavel Urysohn. Sur les classes (L) de M. Fréchet. *Enseignement Mathématique*, 25:77–83, 1926.

Zusammenfassung

Die vorliegende Dissertation bewegt sich auf dem Gebiet der Regularisierung schlecht gestellter inverser Probleme mit einem Fokus auf topologische Aspekte verallgemeinerter Tikhonov-Regularisierung.

Ein inverses Problem besteht nun darin, eine gesuchte Größe, die nur indirekt, zum Beispiel in der Form von Messdaten, die als Ergebnis eines physikalischen Prozesses aus der gesuchten Größe hervorgehen, zu bestimmen.

Mathematisch wird dies als Lösen einer Operatorgleichung $Fx = y$ verstanden, wobei F von der Menge der möglichen gesuchten Größen in die Menge der möglichen Messdaten abbildet, welche beide mit einer Struktur ausgestattet sind, welche die Beschreibung von Nähe beziehungsweise Konvergenz erlaubt.

Ein grundsätzliches Problem, welches beim Lösen solcher Gleichungen auftritt, ist, dass Messdaten prinzipiell fehlerbehaftet sind und zusätzlich inverse Probleme häufig schlecht gestellt sind, das heißt Lösungen im algebraischen Sinn existieren entweder nicht, sind nicht eindeutig oder das Problem ist instabil gegenüber Messfehlern, das heißt aus 'Nähe' im Datenraum lässt sich nicht auf 'Nähe' im Urbildraum schließen.

Der in der Theorie der inversen Probleme verfolgte Ansatz zur Behandlung dieser Problematik erfolgt, jeweils unter Einbeziehung problemspezifischen Vorwissens, zumeist in zwei Schritten: Zuerst wird die Auffassung davon, was man als Lösung versteht, revidiert, was in einem vom algebraischen abweichenden Lösungsbegriff resultiert.

Danach verbleibenden Stabilitätsproblemen wird mit Regularisierung begegnet, deren grundlegende Idee darin besteht, das Problem der Berechnung einer Lösung im Sinne des neuen Lösungsbegriffs durch eine Folge von ähnlichen, stabilen Ersatzproblemen zu ersetzen, deren Lösungen im Falle exakter Lösbarkeit in einer vernünftigen Beziehung zu einer exakten Lösung stehen.

Eine spezielle Variante der Regularisierung ist die verallgemeinerte Tikhonov-Regularisierung, bei der die Folge der Ersatzprobleme in der Minimierung von sogenannten reellwertigen Tikhonovfunktionalen, das heißt Funktionalen der

Form

$$T_{\alpha,y}(x) = \rho(Fx, y) + \alpha R(x)$$

mit einem Parameter α , dem Diskrepanzfunktional ρ , welches 'Nähe' im Datenraum beschreiben soll und einem Regularisierungsfunktional R , bestehen.

Vor diesem Hintergrund bestehen mögliche Aufgaben für die Mathematik nun in der Konstruktion, Untersuchung und Erweiterung von konkreten Verfahren und im Finden von hinreichenden Bedingungen, die deren Korrektheit und wünschenswerte Eigenschaften sicherstellen, aber auch in der Untersuchung mathematischer Objekte und Strukturen auf prinzipielle Nutzbarkeit in diesem Umfeld, sowie im Ausloten der Grenzen herkömmlicher Herangehensweisen. Diese Arbeit legt ihren Fokus vor allem auf die beiden letzten Punkte.

Ein zentraler Punkt in der Theorie inverser Probleme ist die Beschreibung von Nähe und Näherkommen, also Konvergenz in irgendeinem Sinn, im Datenraum. Ausgehend von der These, dass von einem praktischen Gesichtspunkt aus eine Beschreibung von Nähe und Konvergenz durch erweitert reellwertige Funktionale analog zu Metriken wünschenswert ist, Metriken aber zu restriktiv sein können, um Spezifika des konkreten Problems wie zum Beispiel statistisches Verhalten von Messfehlern adäquat zu beschreiben, werden sogenannte Prametriken als Ähnlichkeitsmaß vorgeschlagen. Aus diesen werden analog zur Folgenkonvergenz in metrischen Räumen sequentielle Konvergenzstrukturen, sowie analog zur metrischen Topologie die sogenannte prametrische Topologie konstruiert und deren Eigenschaften untersucht. Insbesondere wird der Frage nachgegangen, wann die von einer Prametrik induzierte sequentielle Konvergenzstruktur topologisch ist. Es wird gezeigt, dass in diesem Fall die prametrische Topologie eine induzierende Topologie und zudem maximal mit dieser Eigenschaft ist.

Weiter zielt die Arbeit auf notwendige Bedingungen an die beteiligten mathematischen Objekte, wie unter anderem Diskrepanz- und Regularisierungsfunktional, sowie die topologischen Strukturen auf Urbild- und Datenraum, ab, um damit ein Regularisierungsverfahren vom Tikhonov-Typ realisieren zu können. Als Vorarbeit dazu werden notwendige Begriffe, wie der eines Regularisierungsoperators, an ein gemischtes Setting aus Topologien und sequentiellen Konvergenzräumen adaptiert und Lösungsbegriffe diskutiert und anschließend notwendige Bedingungen an die Struktur der Lösungsmengen der Ersatzprobleme angegeben.

Des Weiteren wird ein Satz von hinreichenden Standardbedingungen, wie er zum Beispiel in [HKPS07], [Pö08] und [Fle11] verwendet wird, für von Prametriken induzierte sequentielle Konvergenzstrukturen im Datenraum er-

weitere. Neben der Erkenntnis, dass variationelle Standardmethoden weiterhin anwendbar sind, werden notwendige Bedingungen dafür untersucht, dass zwei dieser Voraussetzungen gleichzeitig in einem rein topologischen Rahmen erfüllt sind. Dafür werden sogenannte Bottom-Slice-Topologien konstruiert, die auch für sich selbst als Kandidaten für geeignete Topologien auf dem Datenraum interessant sind.

Abschließend werden ϕ -Bregmanmetriken, eine Klasse von Metriken, die von verallgemeinerten Bregmanabständen abgeleitet sind, näher untersucht.